



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 101709

TO: Celine Qian
Location: CM1/11C10/11E12
Art Unit : 1636
Thursday, August 21, 2003

Case Serial Number: 09/761466

From: Mary Jane Ruhl
Location: Biotech-Chem Library
CM1-6A06
Phone: 605-1155

Maryjane.ruhl@uspto.gov

Search Notes

Examiner Qian,

Here are the results from your search.

Please let me know if you have any questions or need further assistance.

Sincerely,
Mary Jane Ruhl



BEST AVAILABLE COPY


```

;
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,982A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/756,806
; FILING DATE: No. 5981731ember 26, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07111
; FILING DATE: May 31, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,856
; FILING DATE: May 31, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0212
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2510
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
;
US-08-888-982A-42

```

```

Query Match 7.4%; Score 35.2; DB 2; Length 2510;
Best Local Similarity 50.6%; Pred. No. 0.86;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```

```

QY 88 CTTTGTGTACACCTGACCCACAGACGTTCTAGAGAGCCCTTACCGAATAAT 147
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2343 CGTTCTGTGTCACCTGAACAAATGAGTGAAGAGTTGAGAGTGAACAAAGAA 2402
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 AGGAATAATTCCTGACCTGATTGCAAGGCAATCTAATTTTCTTCTCCA 207
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2403 AATAAATGAACATATGTTGCTTATATGTAATAATGAATAAATACCTCTTTT 2462
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 208 GAGCTCTCAAAAAAAAAAAAAAACTTACTTAAAAACAGGATCC 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2463 AGGTGGAATAAAAAAAAAAAAAAAACCC 2510
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 3
US-09-462-261-42
; Sequence 42, Application US/09462261
; Patent No. 6391636
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: Antisense Oligonucleotide
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: Pentium
; OPERATING SYSTEM: Windows 95

```

```

;
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/462,261
; FILING DATE: 01-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/756,806
; FILING DATE: No. 6391636ember 26, 1996
; APPLICATION NUMBER: PCT/US95/07111
; FILING DATE: May 31, 1995
; APPLICATION NUMBER: 08/250,856
; FILING DATE: May 31, 1994
; APPLICATION NUMBER: 08/888,982
; FILING DATE: July 7, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2510
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
;
US-09-462-261-42

```

```

Query Match 7.4%; Score 35.2; DB 4; Length 2510;
Best Local Similarity 50.6%; Pred. No. 0.86;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```

```

QY 88 CTTTGTGTAACTGACCCACAGACTGACGTTCTAGAGAGCCCTTACCGAATAAT 147
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2343 CGTTCTGTGTCACCTGAACAAATGAGTGAAGAGTTGAGAGTGAACAAAGAA 2402
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 148 AGGAATAATTCCTGACCTGATTGCAAGGCAATGCTAATTTTCTTCTCCA 207
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2403 AATAAATGAACATATGTTGCTTATATGTAATAATGAATAAATACCTCTTTT 2462
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 208 GAGCTCTCAAAAAAAAAAAAAAACTTACTTAAAAACAGGATCC 255
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2463 AGGTGGAATAAAAAAAAAAAAAAAACCC 2510
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 4
US-09-506-073-89
; Sequence 89, Application US/09506073
; Patent No. 6410518
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of raf Gene Expression
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/506,073
; EARLIER FILING DATE: 2000-02-18
; EARLIER APPLICATION NUMBER: US 09/143,214
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: PCT/US98/13961
; EARLIER FILING DATE: 1998-07-06
; EARLIER APPLICATION NUMBER: US 08/888,982
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US 08/756,806
; EARLIER FILING DATE: 1996-11-26
; EARLIER APPLICATION NUMBER: PCT/US95/07111
; EARLIER FILING DATE: 1995-05-31
; EARLIER APPLICATION NUMBER: US 08/250,856
; EARLIER FILING DATE: 1994-05-31
; NUMBER OF SEQ ID NOS: 130
; SEQ ID NO 89
; LENGTH: 2510

```



```

1 TYPE: DNA
2 ORGANISM: human sapiens
3 FEATURE:
4 US-09-606 078-879
5
6 query Match 7.44; Score 65.27; 28.4; Length 2710;
7 Best Local Similarity: 59.64; Fold: No. 0.86;
8 Matches: 85; Conserved: 2; Mismatches: 83; Indels: 0; Gaps: 0;
9
10 00 TTTTGTAAATTGTAATGAAATATATACATTTTATATATATGATGTTATGTAATAT 147
11 15 2438 TTTTCTCTTATATGAAATATATATATATATATATATATATATATATATATATAT 2402
12 30 138 ATATAAATAAATATATATATATATATATATATATATATATATATATATATATATAT 207
13 35 2408 ATATAATGAAATATATATATATATATATATATATATATATATATATATATATAT 2462
14 40 2408 GATATCTCAAAAAAAAAAAAAAAAAATATATAAAATAGGATATTTT 275
15 45 2464 AGTTGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTT 2710

```

```

RESULT 5
US-08-700-626-2
SEQUENCE 2 APPLICATION US/97-07-2
Patent No. 674,399
GENERAL INFORMATION:
APPLICANT: ACRYLON 241 G
APPLICANT: ALLIANCE BIOLOGICALS
APPLICANT: TOLL STENO K
TITLE OF INVENTION: NOVEL HUMAN IRI/ACPI-LIKE PROTEIN
NUMBER OF SEQUENCES: 1
CORRESPONDENT ADDRESS:
ADDRESSEE: INCYTE Pharmaceuticals, Inc.
STREET: 4174 Forter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
MOTIF REPEATABLE PERIOD:
MEDIAN TYPE: CISKATTO
COMPILED: IBM COMPILER
OPERATING SYSTEM: DOS
SOFTWARE: FASTSW V08.01.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/98-07-0626
FILING DATE: Filed 6/26/98
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: P0115 US
ELIGIBILITY INFORMATION:
TELEPHONE: 415-855-8555
TELEFAX: 415-845-4176
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 112 Amino acids
TYPE: nucleic acid
STRANDEDNESS: single
DEGREE: 100%
METHOD OF SYNTHESIS: CNA
IMMEDIATE SOURCE:
LIBRARY: pDNAV2.1
JOURNAL: 6/2/98
US-08-700-626-2
Query Match 79% Score 45; DB 1; Length 112;
Host Local Similarity 44.9%; E=0.0; No. 0.71;
Matches 51; Overlap 100 % Mismatches 30; Indels 0; Gaps 0;

```

[illegible]

```

1  RESULT 1 5
2  DS-09-174-400-4
3  SOURCE 5, APPLICANT from US/091474.00
4  ESTD. No. 641561
5  GENERAL INFORMATION:
6  APPLICANT: FAYO, Saverio Carl
7  APPLICANT: HITE, William D.
8  APPLICANT: KIMMOY, Anthony J.
9  APPLICANT: KIMMOY, Rebecca E.
10 APPLICANT: KATASKI, J. Antoni
11 TITLE OF INVENTION: PLANT BRANCH, MAIN AND / WITH RESYNTHETIC INZYMEN
12 FILE REFERENCE: 66 1126
13 CURRENT APPLICATION NUMBER: US/83/174,304
14 CURRENT FILING DATE: 1998-10-15
15 EARLIER APPLICATION NUMBER: 60/064,424
16 EARLIER FILING DATE: 1997 October 28
17 NUMBER OF SEQ ID NOS: 54
18 SOFTWARE: Microsoft Word Version 7.0A
19 SEQ ID NO 3
20 LENGTH: 2074
21 TYPE: DNA
22 ORGANISM: Clivaria max
23 DS-09-174-400-4

```

Only Match	7, 48	Score	4, 4	10h	44	Length	2, 74
Host Local Similarity	7, 34	1004	No.	1, 4			
Matches	44	ConstrValue	0	Mismatches	15	Inserts	0
07	189	1AAATTTTCTTCTCTCAAGCTCT	AAAAAAAAAAAAAAAAAAAA	CTAT	1AAAAA	7, 48	
13	2014	1AAATATTTATCAATTTCTAGAGCTAA	AAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAA	2, 74		

```

RESULT 7
S: 9-460-921b 1b
Sequence ID: Application: US/9946 921b
Patent No.: 607647
GENERAL INFORMATION:
APPLICANT: Leclerc, Joshua Z.
APPLICANT: Budziszewski, Gregory J.
APPLICANT: Patton, Sharon E.
APPLICANT: Wootch, Lynette M.
TITLE OF INVENTION: Herbicide Target Sites and Methods
FILE NUMBER: 09/5 40780A
CURRENT APPLICATION NUMBER: US/99460 921b
CURRENT FILING DATE: 2000-01-11
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patchem Ver. 2.1
SEQ ID No: 1b
LENGTH: 21b
FEATURES:
FEATURE: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_109719
FUNCTION: (1) : (71b)
OTHER INFORMATION: a, b, c, d, e
US-99-460 921b 1b

```

[illegible]

Db 161 COTTATTAATTTTATTGATAGCCAAAAA 214
US-09-149-476-191
; Sequence 191, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22

```
1 EARLIER APPLICATION NUMBER: 60/057,761
2 EARLIER FILING DATE: 1997-08-22
3 EARLIER APPLICATION NUMBER: 60/047,535
4 EARLIER FILING DATE: 1997-05-24
5 EARLIER APPLICATION NUMBER: 60/047,599
6 EARLIER FILING DATE: 1997-05-24
7 EARLIER APPLICATION NUMBER: 60/047,598
8 EARLIER FILING DATE: 1997-05-24
9 EARLIER APPLICATION NUMBER: 60/047,585
10 EARLIER FILING DATE: 1997-05-24
11 EARLIER APPLICATION NUMBER: 60/047,586
12 EARLIER FILING DATE: 1997-05-24
13 EARLIER APPLICATION NUMBER: 60/047,590
14 EARLIER FILING DATE: 1997-05-23
15 EARLIER APPLICATION NUMBER: 60/047,594
16 EARLIER FILING DATE: 1997-05-23
17 EARLIER APPLICATION NUMBER: 60/047,589
18 EARLIER FILING DATE: 1997-05-23
19 EARLIER APPLICATION NUMBER: 60/047,593
20 EARLIER FILING DATE: 1997-05-23
21 EARLIER APPLICATION NUMBER: 60/047,614
22 EARLIER FILING DATE: 1997-05-23
23 EARLIER APPLICATION NUMBER: 60/043,578
24 EARLIER FILING DATE: 1997-04-11
25 EARLIER APPLICATION NUMBER: 60/043,576
26 EARLIER FILING DATE: 1997-04-11
27 EARLIER APPLICATION NUMBER: 60/047,501
28 EARLIER FILING DATE: 1997-05-22
29 EARLIER APPLICATION NUMBER: 60/043,670
30 EARLIER FILING DATE: 1997-04-11
31 EARLIER APPLICATION NUMBER: 60/056,632
32 EARLIER FILING DATE: 1997-08-22
33 EARLIER APPLICATION NUMBER: 60/056,664
34 EARLIER FILING DATE: 1997-08-22
35 EARLIER APPLICATION NUMBER: 60/056,876
36 EARLIER FILING DATE: 1997-08-22
37 EARLIER APPLICATION NUMBER: 60/056,881
38 EARLIER FILING DATE: 1997-08-22
39 EARLIER APPLICATION NUMBER: 60/056,909
40 EARLIER FILING DATE: 1997-08-22
41 EARLIER APPLICATION NUMBER: 60/056,875
42 EARLIER FILING DATE: 1997-08-22
43 EARLIER APPLICATION NUMBER: 60/056,862
44 EARLIER FILING DATE: 1997-08-22
45 EARLIER APPLICATION NUMBER: 60/056,867
46 EARLIER FILING DATE: 1997-08-22
47 EARLIER APPLICATION NUMBER: 60/056,908
48 EARLIER FILING DATE: 1997-08-22
49 EARLIER APPLICATION NUMBER: 60/048,964
50 EARLIER FILING DATE: 1997-08-06
51 EARLIER APPLICATION NUMBER: 60/057,650
52 EARLIER FILING DATE: 1997-03-05
53 EARLIER APPLICATION NUMBER: 60/056,884
54 EARLIER FILING DATE: 1997-08-22
55 EARLIER APPLICATION NUMBER: 60/057,669
56 EARLIER FILING DATE: 1997-09-05
57 EARLIER APPLICATION NUMBER: 60/049,610
58 EARLIER FILING DATE: 1997-06-13
59 EARLIER APPLICATION NUMBER: 60/061,060
60 EARLIER FILING DATE: 1997-10-02

Query Match: 7.98; Score 44.6; DB 4; Length 2779;
Best Local Similarity: 61.44; Pseq. No. 2.0;
Matches 54; Conservative 9; Mismatches 34; Indels 0; Gaps 0;

153 GGTACTTGATTTGGAGAGTAATCTTAATTTTCTTCCTCAGAGCTCAAAAAA 222
154 1111111111111111111111111111111111111111111111111
155 2691 GGTCCAGTGAAATCTTATTTCTCAATATATATTTGATTAAGTCAAAAAA 2750
156 223 AAAAAAAGCTTACTTAAAAAAGG 250
157 1111111111111111111111111111111111111111111111111
158 2751 AAAAAAAGCTTACTTAAAAAAGG 2778
```

```
858011 9
US-09-227-457-40
1 Sequence 40, Application US/09227457
2 Patent No. 6342581
3 GENERAL INFORMATION:
4 APPLICANT: Fischer et al.
5 TITLE OF INVENTION: 123 Human Secreted Proteins
6 FILE REFERENCE: P2010P1
7 CURRENT FILING DATE: 1999-01-08
8 EARLIER APPLICATION NUMBER: PCT/US98/13664
9 EARLIER FILING DATE: 1998-07-07
10 EARLIER APPLICATION NUMBER: 60/051,926
11 EARLIER FILING DATE: 1997-07-08
12 EARLIER APPLICATION NUMBER: 60/052,793
13 EARLIER FILING DATE: 1997-07-08
14 EARLIER APPLICATION NUMBER: 60/051,925
15 EARLIER FILING DATE: 1997-07-08
16 EARLIER APPLICATION NUMBER: 60/051,929
17 EARLIER FILING DATE: 1997-07-08
18 EARLIER APPLICATION NUMBER: 60/052,803
19 EARLIER FILING DATE: 1997-07-08
20 EARLIER APPLICATION NUMBER: 60/052,732
21 EARLIER FILING DATE: 1997-07-08
22 EARLIER APPLICATION NUMBER: 60/051,931
23 EARLIER FILING DATE: 1997-07-08
24 EARLIER APPLICATION NUMBER: 60/051,932
25 EARLIER FILING DATE: 1997-07-08
26 EARLIER APPLICATION NUMBER: 60/051,916
27 EARLIER FILING DATE: 1997-07-08
28 EARLIER APPLICATION NUMBER: 60/051,930
29 EARLIER FILING DATE: 1997-07-08
30 EARLIER APPLICATION NUMBER: 60/051,918
31 EARLIER FILING DATE: 1997-07-08
32 EARLIER APPLICATION NUMBER: 60/051,920
33 EARLIER FILING DATE: 1997-07-08
34 EARLIER APPLICATION NUMBER: 60/052,733
35 EARLIER FILING DATE: 1997-07-08
36 EARLIER APPLICATION NUMBER: 60/052,795
37 EARLIER FILING DATE: 1997-07-08
38 EARLIER APPLICATION NUMBER: 60/051,919
39 EARLIER FILING DATE: 1997-07-08
40 EARLIER APPLICATION NUMBER: 60/051,928
41 EARLIER FILING DATE: 1997-07-08
42 EARLIER APPLICATION NUMBER: 60/055,722
43 EARLIER FILING DATE: 1997-08-18
44 EARLIER APPLICATION NUMBER: 60/055,723
45 EARLIER FILING DATE: 1997-08-18
46 EARLIER APPLICATION NUMBER: 60/055,948
47 EARLIER FILING DATE: 1997-08-18
48 EARLIER APPLICATION NUMBER: 60/055,949
49 EARLIER FILING DATE: 1997-08-18
50 EARLIER APPLICATION NUMBER: 60/055,953
51 EARLIER FILING DATE: 1997-08-18
52 EARLIER APPLICATION NUMBER: 60/055,950
53 EARLIER FILING DATE: 1997-08-18
54 EARLIER APPLICATION NUMBER: 60/055,947
55 EARLIER FILING DATE: 1997-08-18
56 EARLIER APPLICATION NUMBER: 60/055,964
57 EARLIER FILING DATE: 1997-08-18
58 EARLIER APPLICATION NUMBER: 60/056,960
59 EARLIER FILING DATE: 1997-08-18
60 EARLIER APPLICATION NUMBER: 60/056,983
61 EARLIER FILING DATE: 1997-08-18
62 EARLIER APPLICATION NUMBER: 60/056,984
63 EARLIER FILING DATE: 1997-08-18
64 EARLIER APPLICATION NUMBER: 60/056,983
65 EARLIER FILING DATE: 1997-08-18
66 EARLIER APPLICATION NUMBER: 60/056,983
67 EARLIER FILING DATE: 1997-08-18
68 EARLIER APPLICATION NUMBER: 60/056,983
69 EARLIER FILING DATE: 1997-08-18
70 EARLIER APPLICATION NUMBER: 60/056,983
71 EARLIER FILING DATE: 1997-08-18
72 EARLIER APPLICATION NUMBER: 60/056,983
73 EARLIER FILING DATE: 1997-08-18
74 EARLIER APPLICATION NUMBER: 60/056,983
75 EARLIER FILING DATE: 1997-08-18
76 EARLIER APPLICATION NUMBER: 60/056,983
77 EARLIER FILING DATE: 1997-08-18
78 EARLIER APPLICATION NUMBER: 60/056,983
79 EARLIER FILING DATE: 1997-08-18
80 EARLIER APPLICATION NUMBER: 60/056,983
81 EARLIER FILING DATE: 1997-08-18
82 EARLIER APPLICATION NUMBER: 60/056,983
83 EARLIER FILING DATE: 1997-08-18
84 EARLIER APPLICATION NUMBER: 60/056,983
85 EARLIER FILING DATE: 1997-08-18
86 EARLIER APPLICATION NUMBER: 60/056,983
87 EARLIER FILING DATE: 1997-08-18
88 EARLIER APPLICATION NUMBER: 60/056,983
89 EARLIER FILING DATE: 1997-08-18
90 EARLIER APPLICATION NUMBER: 60/056,983
91 EARLIER FILING DATE: 1997-08-18
92 EARLIER APPLICATION NUMBER: 60/056,983
93 EARLIER FILING DATE: 1997-08-18
94 EARLIER APPLICATION NUMBER: 60/056,983
95 EARLIER FILING DATE: 1997-08-18
96 EARLIER APPLICATION NUMBER: 60/056,983
97 EARLIER FILING DATE: 1997-08-18
98 EARLIER APPLICATION NUMBER: 60/056,983
99 EARLIER FILING DATE: 1997-08-18
100 EARLIER APPLICATION NUMBER: 60/056,983
```

```

: EARLIER APPLICATION NUMBER: 60/058,660
: EARLIER FILING DATE: 1997-09-12
: EARLIER APPLICATION NUMBER: 60/058,661
: EARLIER FILING DATE: 1997-09-12
: NUMBER OF SEQ ID NOS: 672
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 40
: LENGTH: 1129
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1053)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-40

Query Match
Best Local Similarity 6.9%; Score 33.2; DB 4; Length 1129;
Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 170 TGAATTGACAGGCAATGCTATTTTCTTCTCCAGAGCTCAAAAAAAAAAAAAA 229
Db 1016 TGGTGAGCAATGACACATAATTTTCATTTTCCNAAAAAAAAAAAAAAAAA 1075
QY 230 AAAAAGCTTACTAAAAACA 248
Db 1076 AAAAAAAAAAAAAAAAAA 1094

RESULT 10
US-09-489-847-101
: Sequence 101, Application US/09489847
: Patent No. 6476195
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: 98 Human Secreted Proteins
: FILE REFERENCE: P2031P1
: CURRENT APPLICATION NUMBER: US/09-489,847
: EARLIER FILING DATE: 2000-01-24
: EARLIER APPLICATION NUMBER: PCT/US99/17130
: EARLIER FILING DATE: 1999-07-29
: EARLIER APPLICATION NUMBER: 60/094,657
: EARLIER FILING DATE: 1998-07-30
: EARLIER APPLICATION NUMBER: 60/095,486
: EARLIER FILING DATE: 1998-08-05
: EARLIER APPLICATION NUMBER: 60/096,319
: EARLIER FILING DATE: 1998-08-12
: EARLIER APPLICATION NUMBER: 60/095,454
: EARLIER FILING DATE: 1998-08-06
: EARLIER APPLICATION NUMBER: 60/095,455
: EARLIER FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 376
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 101
: LENGTH: 1248
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-489-847-101

Query Match
Best Local Similarity 6.9%; Score 33.2; DB 4; Length 1248;
Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 195 TTTTCTTCTCAGAGCTCTCAAAAAAAAAAAAAAAGCTTACTAAAAACA 248
Db 1141 TTATATCTCTCAAAAGCTCTCAAAAAAAAAAAAAAAGCTTACTAAAAACA 1194

RESULT 11
US-09-702-705-1801
: Sequence 1801, Application US/09702705
: Patent No. 6504010
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION:
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER:
: EARLIER FILING DATE:
: NUMBER OF SEQ ID NOS:
: SOFTWARE:
: SEQ ID NO:
: LENGTH:
: TYPE:
: ORGANISM:
US-09-702-705-1801
```

```

: APPLICANT: Wang, Tonglong
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Lodes, Michael A.
: APPLICANT: Fanger, Gary
: APPLICANT: Vedvick, Tom
: APPLICANT: Carter, Darrick
: APPLICANT: Retter, Marc
: APPLICANT: Mannion, Jane
: APPLICANT: Fan, Liqun
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.478C14
: CURRENT APPLICATION NUMBER: US/09/702,705
: EARLIER FILING DATE: 2000-10-30
: NUMBER OF SEQ ID NOS: 1833
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1801
: LENGTH: 4086
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-702-705-1801

Query Match
Best Local Similarity 6.9%; Score 33.2; DB 4; Length 4086;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 123 CTAGAGAGCCGCCCTTACCCGAATAATGAAATATCCCTTGCACCTTGATTGCAAGG 182
Db 3910 CTGGGCTGGCTGATTCACGATGACGAGATGTCATAATATACAGATAGAGATGATGCT 3969
QY 183 CATGCTAATTTTCTTCTCTCCAGAGCTCTCAAAAAAAAAAAAAAAGCTTACTA 242
Db 3970 ACCTATATAAGACTTTTTCCTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4029
QY 243 AAAACA 248
Db 4030 AAAAAA 4035

RESULT 12
US-09-736-457-1801
: Sequence 1801, Application US/09736457
: Patent No. 6509448
: GENERAL INFORMATION:
: APPLICANT: Wang, Tonglong
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Lodes, Michael A.
: APPLICANT: Fanger, Gary
: APPLICANT: Vedvick, Tom
: APPLICANT: Carter, Darrick
: APPLICANT: Retter, Marc
: APPLICANT: Mannion, Jane
: APPLICANT: Fan, Liqun
: APPLICANT: Wang, Aijun
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.478C15
: CURRENT APPLICATION NUMBER: US/09/736,457
: EARLIER FILING DATE: 2000-12-13
: NUMBER OF SEQ ID NOS: 1864
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1801
: LENGTH: 4086
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-736-457-1801

Query Match
Best Local Similarity 6.9%; Score 33.2; DB 4; Length 4086;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 123 CTAGAGAGCCGCCCTTACCCGAATAATGAAATATCCCTTGCACCTTGATTGCAAGG 182
Db 3910 CTGGGCTGGCTGATTCACGATGACGAGATGTCATAATATACAGATAGAGATGATGCT 3969
QY 183 CATGCTAATTTTCTTCTCTCCAGAGCTCTCAAAAAAAAAAAAAAAGCTTACTA 242
Db 3970 ACCTATATAAGACTTTTTCCTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 4029
QY 243 AAAACA 248
Db 4030 AAAAAA 4035
```



```

PCT-US94-14073-1
? Sequence 1, Application PC/TUS9414073
? GENERAL INFORMATION:
? APPLICANT: THOMAS, Wayne R., CHUA, Kaw-Yan, ROGERS, Bruce L., and
? APPLICANT: KUO, Mei-chang
? TITLE OF INVENTION: NUCLEIC ACIDS ENCODING A HOUSE DUST MITE
? TITLE OF INVENTION: ALLERGEN, DER P IIT, AND USES THEREFOR
? NUMBER OF SEQUENCES: 18
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Lahive & Cockfield
? STREET: 60 State Street suite 510
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII TEXT
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US94/14073
? FILING DATE: 08-DEC-1993
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? CLASSIFICATION:
? ATTORNEY/AGENT INFORMATION:
? NAME: Mandragoras, Amy E.
? REGISTRATION NUMBER: 36,207
? REFERENCE/DOCKET NUMBER: IPC-072 (IMI-041)
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1059 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 63..848
? FEATURE:
? NAME/KEY: mat_peptide
? LOCATION: 150..848
PCT-US94-14073-1

Query Match 6.9%; Score 33; DB 5; Length 1059;
Best Local Similarity 57.1%; Pred.No. 2.7;
Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 144 AATAATGCAATTAATCTTGCACCTTGATTTGCAAGGCAATGCTAATTTTTCCTTC 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 951 AATATTCGAGCTGCAAAACCAACCAATCATGATGTATAGGAGTTGTGTTGTTTCAATT 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 204 TCCAGAGCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAACCCTTACATAAACA 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1011 TAAAAAATCCAAAAATTAATTAATTAATTAATCAATTCACAAAAA 1055
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: August 20, 2003, 14:32:46
Job time : 51.3482 secs

PT Induction, e.g. for optimizing cardiomyocyte induction
 XX
 PS Claim 21; Fig 5c; 66pp; English.
 XX
 CC The sequence represents the coding sequence of cardiac enhancer
 CC hcsx/Nkx2.5 homology domain B. The nucleic acid is useful for
 CC specifically expressing a gene in a cardiac cell, as an earlier marker
 CC of cardiomyocyte induction, e.g. for optimizing cardiomyocyte induction.
 CC Genes expressed in the cardiac cell-specific manner are useful for the
 CC targeted expression of genes encoding therapeutic proteins for the
 CC treatment of damaged heart tissue. Cardiac specific enhancer elements may
 CC be used for gene therapy.
 CC
 XX
 SQ Sequence 478 BP; 138 A; 129 C; 89 G; 122 T; 0 other;
 Query Match 100.0%; Score 478; DB 22; Length 478;
 Best Local Similarity 100.0%; Pred. No. 2.2e-122;
 Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGAGAAATCATTTACCCGATTCACAAAGAGCATAGAGAGTGAACAGTCACTGATCTTGT 60
 DB 1 AGAGAAATCATTTACCCGATTCACAAAGAGCATAGAGAGTGAACAGTCACTGATCTTGT 60
 QY 61 CAAATAGGAGAGATTTTTTCTTCCCTTTTGTACACCTGACCCACAGAGACTGACAG 120
 DB 61 CAAATAGGAGAGATTTTTTCTTCCCTTTTGTACACCTGACCCACAGAGACTGACAG 120
 QY 121 TTCTAGAGAGCCCCCTTACCCGAAATAGAAATTAATCCTTGGCACCCTGATTTGCAG 180
 DB 121 TTCTAGAGAGCCCCCTTACCCGAAATAGAAATTAATCCTTGGCACCCTGATTTGCAG 180
 QY 181 GGCATGCTAATTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTAAACCTTAC 240
 DB 181 GGCATGCTAATTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTAAACCTTAC 240
 QY 241 TAAATAAGAGGATCCCGGATGTACCTCGATGTCCCCATTAACGGTAATTTTCAGGC 300
 DB 241 TAAATAAGAGGATCCCGGATGTACCTCGATGTCCCCATTAACGGTAATTTTCAGGC 300
 QY 301 GTCCGCTACACTAATCTTTCAAACTGTCATCGGAGCGGCTGGCCAGAGATTCACCTT 360
 DB 301 GTCCGCTACACTAATCTTTCAAACTGTCATCGGAGCGGCTGGCCAGAGATTCACCTT 360
 QY 361 AACAGCGCTCCAGAGACCCCTGTCGAGCTCTTTTCAGGAGACATTTAATTGAATGG 420
 DB 361 AACAGCGCTCCAGAGACCCCTGTCGAGCTCTTTTCAGGAGACATTTAATTGAATGG 420
 QY 421 ATGTGGCTGTTTGGCAGAGCTCAACGCTTGGGAGATAGGCATCTCTCCAAAGCAGAC 478
 DB 421 ATGTGGCTGTTTGGCAGAGCTCAACGCTTGGGAGATAGGCATCTCTCCAAAGCAGAC 478
 DB
 RESULT 2
 AAS09962
 ID AAS09962 standard; DNA; 6751 BP.
 XX
 AC AAS09962;
 XX
 DT 24-Oct-2001 (first entry)
 XX
 DE Genomic DNA #2 encoding human Csx/Nkx2.5.
 XX
 KW Csx/Nkx2.5; cardiac enhancer; cardiac cell; cardiomyocyte induction;
 KW therapeutic; heart tissue; gene therapy; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN W0200151006-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001WO-US01511.
 XX

PR 14-JAN-2000; 2000US-0176419.
 XX
 PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
 XX
 PI Lee IW, Izumo S;
 XX
 DR WPI; 2001-451809/48.
 XX
 PT New cardiac specific cell enhancer elements, useful for specifically
 PT expressing gene in cardiac cell, as earlier marker of cardiomyocyte
 PT induction, e.g. for optimizing cardiomyocyte induction
 CC
 XX
 PS Claim 29; Fig 4b; 66pp; English.
 XX
 CC The sequence represents the genomic sequence #2 of human Csx/Nkx2.5. The
 CC nucleic acid is useful for specifically expressing a gene in a cardiac
 CC cell, as an earlier marker of cardiomyocyte induction, e.g. for
 CC optimizing cardiomyocyte induction. Genes expressed in the cardiac cell-
 CC specific manner are useful for the targeted expression of genes encoding
 CC therapeutic proteins for the treatment of damaged heart tissue. Cardiac
 CC specific enhancer elements may be used for gene therapy.
 CC
 XX
 SQ Sequence 6751 BP; 1459 A; 2008 C; 1753 G; 1490 T; 41 other;
 Query Match 100.0%; Score 478; DB 22; Length 6751;
 Best Local Similarity 100.0%; Pred. No. 6e-122;
 Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGAGAAATCATTTACCCGATTCACAAAGAGCATAGAGAGTGAACAGTCACTGATCTTGT 60
 DB 201 AGAGAAATCATTTACCCGATTCACAAAGAGCATAGAGAGTGAACAGTCACTGATCTTGT 260
 QY 61 CAAATAGGAGAGATTTTTTCTTCCCTTTTGTACACCTGACCCACAGAGACTGACAG 120
 DB 261 CAAATAGGAGAGATTTTTTCTTCCCTTTTGTACACCTGACCCACAGAGACTGACAG 320
 QY 121 TTCTAGAGAGCCCCCTTACCCGAAATAGAAATTAATCCTTGGCACCCTGATTTGCAG 180
 DB 321 TTCTAGAGAGCCCCCTTACCCGAAATAGAAATTAATCCTTGGCACCCTGATTTGCAG 380
 QY 381 GGCATGCTAATTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTAAACCTTAC 240
 DB 381 GGCATGCTAATTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTAAACCTTAC 440
 QY 241 TAAATAAGAGGATCCCGGATGTACCTCGATGTCCCCATTAACGGTAATTTTCAGGC 300
 DB 441 TAAATAAGAGGATCCCGGATGTACCTCGATGTCCCCATTAACGGTAATTTTCAGGC 500
 QY 301 GTCCGCTACACTAATCTTTCAAACTGTCATCGGAGCGGCTGGCCAGAGATTCACCTT 360
 DB 501 GTCCGCTACACTAATCTTTCAAACTGTCATCGGAGCGGCTGGCCAGAGATTCACCTT 560
 QY 361 AACAGCGCTCCAGAGACCCCTGTCGAGCTCTTTTCAGGAGACATTTAATTGAATGG 420
 DB 561 AACAGCGCTCCAGAGACCCCTGTCGAGCTCTTTTCAGGAGACATTTAATTGAATGG 620
 QY 421 ATGTGGCTGTTTGGCAGAGCTCAACGCTTGGGAGATAGGCATCTCTCCAAAGCAGAC 478
 DB 621 ATGTGGCTGTTTGGCAGAGCTCAACGCTTGGGAGATAGGCATCTCTCCAAAGCAGAC 678
 DB
 RESULT 3
 AAC59534
 ID AAC59534 standard; cDNA; 1086 BP.
 XX
 AC AAC59534;
 XX
 DT 15-FEB-2001 (first entry)
 XX
 DE Human secreted protein cDNA sequence #28.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;

PR	14-SEP-2000	2000US-02332398
PR	14-SEP-2000	2000US-02332397
PR	14-SEP-2000	2000US-02332399
PR	14-SEP-2000	2000US-02324401
PR	14-SEP-2000	2000US-02324400
PR	14-SEP-2000	2000US-02330664
PR	14-SEP-2000	2000US-02330663
PR	14-SEP-2000	2000US-02330665
PR	21-SEP-2000	2000US-02344223
PR	21-SEP-2000	2000US-02344224
PR	25-SEP-2000	2000US-02343997
PR	25-SEP-2000	2000US-02343998
PR	26-SEP-2000	2000US-02354584
PR	27-SEP-2000	2000US-02358534
PR	27-SEP-2000	2000US-02358536
PR	29-SEP-2000	2000US-02363627
PR	29-SEP-2000	2000US-02363628
PR	29-SEP-2000	2000US-02363659
PR	29-SEP-2000	2000US-02363670
PR	02-OCT-2000	2000US-02373037
PR	02-OCT-2000	2000US-02373038
PR	02-OCT-2000	2000US-02373039
PR	02-OCT-2000	2000US-02370740
PR	13-OCT-2000	2000US-02393935
PR	13-OCT-2000	2000US-02393937
PR	20-OCT-2000	2000US-02341821
PR	20-OCT-2000	2000US-02341820
PR	20-OCT-2000	2000US-02341819
PR	20-OCT-2000	2000US-02341818
PR	20-OCT-2000	2000US-02341815
PR	20-OCT-2000	2000US-02417187
PR	20-OCT-2000	2000US-02416775
PR	20-OCT-2000	2000US-02464676
PR	08-NOV-2000	2000US-02464675
PR	08-NOV-2000	2000US-02464677
PR	08-NOV-2000	2000US-02464678
PR	08-NOV-2000	2000US-02465528
PR	08-NOV-2000	2000US-02465528
PR	08-NOV-2000	2000US-02465639
PR	08-NOV-2000	2000US-02466110
PR	08-NOV-2000	2000US-02466111
PR	08-NOV-2000	2000US-02466132
PR	17-NOV-2000	2000US-02492037
PR	17-NOV-2000	2000US-02492038
PR	17-NOV-2000	2000US-02492039
PR	17-NOV-2000	2000US-02492100
PR	17-NOV-2000	2000US-02492111
PR	17-NOV-2000	2000US-02492112
PR	17-NOV-2000	2000US-02492913
PR	17-NOV-2000	2000US-02492914
PR	17-NOV-2000	2000US-02492945
PR	17-NOV-2000	2000US-02492946
PR	17-NOV-2000	2000US-02492965
PR	17-NOV-2000	2000US-02492967
PR	17-NOV-2000	2000US-02493927
PR	17-NOV-2000	2000US-02493928
PR	17-NOV-2000	2000US-02493930
PR	01-DEC-2000	2000US-02501660
PR	01-DEC-2000	2000US-02501661

PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-457723/49.
 XX
 PT Isolated polypeptide for treating, preventing and/or prognosing
 PT respiratory disorders related to the lung including lung cancers and
 PT also for testing and detection e.g. diagnosis -
 XX
 PS Claim 1; SEQ ID No 363; 507pp; English.

CC Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC and their associated polynucleotides are useful in the diagnosis,
CC treatment and prevention of various types of disorders in e.g. humans,
CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC pathological condition can be determined by detecting the presence or
CC absence of a mutation in a lung antigen polynucleotide. The treatable
CC disorders include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://www.int/pub/published_pct_sequences](http://www.int/pub/published_pct_sequences).

Query Match	8.2%;	Score 39;	DB 22;	Length 3273;
Best Local Similarity	68.4%;	Pred. No. 1.3;		
Matches 54;	Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;

Oy 177 CAAGGGCAATGCTAATTTTTCCTCCTCCAGACCTCTCAAAAAAAAAAAAAAAC 236
| | | | |
Dd 2059 CATTGACCATGTAAATTTATTCCTCTTACCCCTGTTAAAAAAACTTAAAAAAC 2118
| | | | |

OY	237	TTACTAAAAACAGGGATCC	255
DB	2119	TAAAAAATAAAAGAATCC	2137

RESULT	5
AAS30098	
ID	AAS30098 standard; DNA; 3276 BP.

AC AAS300987

DT 21-NOV-2001 (first entry)

Human lung antigen genomic DNA #168.

Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW

PR	06-DEC-2000	2000US-0251479.	
PR	08-DEC-2000	2000US-0251856.	
PR	08-DEC-2000	2000US-0251868.	
PR	08-DEC-2000	2000US-0251869.	
PR	08-DEC-2000	2000US-0251989.	
PR	08-DEC-2000	2000US-0251990.	
PR	11-DEC-2000	2000US-0254097.	
PR	05-JAN-2001	2001US-0259678.	
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA, Barash SC, Ruben SW;		
DR	WPI: 2001-457723/49.		
XX			
PT	Isolated polypeptide for treating, preventing and/or prognosing		
PR	respiratory disorders related to the lung including lung cancers and		
PT	also for testing and detection e.g. diagnosis -		
XX			
PS	Claim 1; SEQ ID No 362; 507pp; English.		
XX			
CC	Sequences AAS293931-AAS30164 represent genomic DNA molecules, which encode		
CC	the lung antigen polypeptides of the invention. Lung antigen polypeptides		
CC	and their associated polynucleotides are useful in the diagnosis,		
CC	treatment and prevention of various types of disorders in e.g. humans,		
CC	mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A		
CC	pathological condition can be determined by detecting the presence or		
CC	absence of a mutation in a lung antigen polynucleotide. The treatable		
CC	disorders include autoimmune diseases such as rheumatoid arthritis,		
CC	hyperproliferative disorders such as neoplasms of the breast or liver,		
CC	cardiovascular disorders such as cardiac arrest, cerebrovascular		
CC	disorders such as cerebral ischaemia, nervous system disorders such as		
CC	Alzheimer's disease, infections caused by bacteria, viruses and fungi,		
CC	ocular disorders such as corneal infection, endocrine disorders such as		
CC	premature labour and infertility, gastrointestinal disorders such as		
CC	Crohn's disease, renal disorders such as glomerulonephritis and		
CC	respiratory disorders such as asthma and pleurisy. The polypeptides can		
CC	also be used to aid wound healing, to prevent skin aging due to sunburn,		
CC	to maintain organs before transplantation, to regenerate tissues and in		
CC	chemotaxis. The polypeptides can also be used as a food additive or		
CC	preservative to increase or decrease storage capabilities.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
Query Match	8.2%;	Score 39;	DB 22; Length 3276;
Best Local Similarity	68.4%;	Pred. No. 1.3;	
Matches 54;	Conservative 0;	Mismatches 25;	Indels 0; Gaps 0
Oy	177 CAAGGCGCAATGCTATTTTCTTCCTCCCAAGCTCTCAAAAAAAAAAAAAAAAAAAC	236	
Db	2059 CATGGACATGTAATTTTATCTCTTCCTACCTGTTAAAAAACTTAAAAAAAC	211	
Oy	237 TTACTTAAACAGAGGATCC	255	
Db	2119 TAAAAAATAAAAAAGTCC	2137	
RESULT 6			
AA193585/c			
AA193585 standard; cDNA; 397 BP.			
AA193585;			
06-NOV-2001 (first entry)			
Human polynucleotide SEQ ID NO 13645.			
Human: cytokine; cell proliferation; cell differentiation; gene therapy;			
vaccine; peptide therapy; stem cell growth factor; haematopoiesis;			
tissue growth factor; immunomodulatory; cancer; leukaemia;			
nervous system disorders; arthritis; inflammation; ss.			

PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241825.
PR 01-NOV-2000; 2000US-0241826.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0234097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-457723/49.
DR P-PSDB; AAU18622.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT respiratory disorders related to the lung including lung cancers and
PT also for testing and detection e.g. diagnosis -
XX
XX Claim 1; SEQ ID No 81; 507pp; English.
CC Sequences AAS29839-AAS29930 represent cDNA molecules, which encode the
CC lung antigen polypeptides of the invention. Lung antigen polypeptides and
CC their associated polynucleotides are useful in the diagnosis, treatment
CC and prevention of various types of disorders in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological
CC condition can be determined by detecting the presence or absence of a
CC mutation in a lung antigen polynucleotide. The treatable disorders
CC include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

Query Match 8.0%; Score 38.4; DB 22; Length 498;
Best Local Similarity 67.5%; Pred. No. 0.97;
Matches 54; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 177 CAAGGCAATGCTAATTTTCTTCCTCCACAGCTCTCAAAAAAAAAAAAAAAAAACC 236
DB 412 CATGGACAAATGTAATTTTATCTCTACCTGTTAAAAAACTAAAAAAACC 471

OY 237 TTACTAAAAACAGGATGCC 256
DB 472 TAAAAAATAAAAAAAAAAGTCC 491

RESULT 9
AAF94842/c
ID AAF94842 strand: cDNA; 396 BP.
XX
AC AAF94842;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 33.
XX
KW Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200118046-A2.
XX
PD 15-MAR-2001.
XX

XX
 67 Testing primers and probes for analysing diseases associated with
 68 cytosine methylation state e.g. arthritis, cancer, aging,
 69 arteriosclerosis comprising fragments of chemically modified genes
 70 associated with cell cycle
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100

Claim 1: SEQ ID No 145; 26pp; English.

XX
 101 Sequences AAS452/6-AAS452/7 represent chemically pretreated genomic DNA
 102 molecules associated with the cell cycle and specific PCR primers of the
 103 invention. The sequences are useful for detecting the methylation state
 104 of all CpG dinucleotides in a sequence and therefore for analysing
 105 associated diseases. By analysing cytosine methylations in the pretreated
 106 DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 107 of existing diseases or the predisposition to specific diseases can be
 108 ascertained. The parameters may be compared to another set of genetic
 109 and/or epigenetic parameters, the differences serving as basis for
 110 diagnosis and/or prognosis events which are disadvantageous to patients.
 111 The sequences of the invention are useful for the diagnosis and therapy
 112 of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 113 aging, glomerular disease, low body disease, arthritis,
 114 arteriosclerosis, solid tumours and cancers.
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908
 909
 910
 911
 912
 913
 914
 915
 916
 917
 918
 919
 920
 921
 922
 923
 924
 925
 926
 927
 928
 929
 930
 931
 932
 933
 934
 935
 936
 937
 938
 939
 940
 941
 942
 943
 944
 945
 946
 947
 948
 949
 950
 951
 952
 953
 954
 955
 956
 957
 958
 959
 960
 961
 962
 963
 964
 965
 966
 967
 968
 969
 970
 971
 972
 973
 974
 975
 976
 977
 978
 979
 980
 981
 982
 983
 984
 985
 986
 987
 988
 989
 990
 991
 992
 993
 994
 995
 996
 997
 998
 999
 1000

XX
 1001 Sequence 11047 BF: 3247 A: 218 C: 2358 G: 5204 T: 0 Other:
 1002
 1003
 1004
 1005
 1006
 1007
 1008
 1009
 1010
 1011
 1012
 1013
 1014
 1015
 1016
 1017
 1018
 1019
 1020
 1021
 1022
 1023
 1024
 1025
 1026
 1027
 1028
 1029
 1030
 1031
 1032
 1033
 1034
 1035
 1036
 1037
 1038
 1039
 1040
 1041
 1042
 1043
 1044
 1045
 1046
 1047
 1048
 1049
 1050
 1051
 1052
 1053
 1054
 1055
 1056
 1057
 1058
 1059
 1060
 1061
 1062
 1063
 1064
 1065
 1066
 1067
 1068
 1069
 1070
 1071
 1072
 1073
 1074
 1075
 1076
 1077
 1078
 1079
 1080
 1081
 1082
 1083
 1084
 1085
 1086
 1087
 1088
 1089
 1090
 1091
 1092
 1093
 1094
 1095
 1096
 1097
 1098
 1099
 1100
 1101
 1102
 1103
 1104
 1105
 1106
 1107
 1108
 1109
 1110
 1111
 1112
 1113
 1114
 1115
 1116
 1117
 1118
 1119
 1120
 1121
 1122
 1123
 1124
 1125
 1126
 1127
 1128
 1129
 1130
 1131
 1132
 1133
 1134
 1135
 1136
 1137
 1138
 1139
 1140
 1141
 1142
 1143
 1144
 1145
 1146
 1147
 1148
 1149
 1150
 1151
 1152
 1153
 1154
 1155
 1156
 1157
 1158
 1159
 1160
 1161
 1162
 1163
 1164
 1165
 1166
 1167
 1168
 1169
 1170
 1171
 1172
 1173
 1174
 1175
 1176
 1177
 1178
 1179
 1180
 1181
 1182
 1183
 1184
 1185
 1186
 1187
 1188
 1189
 1190
 1191
 1192
 1193
 1194
 1195
 1196
 1197
 1198
 1199
 1200
 1201
 1202
 1203
 1204
 1205
 1206
 1207
 1208
 1209
 1210
 1211
 1212
 1213
 1214
 1215
 1216
 1217
 1218
 1219
 1220
 1221
 1222
 1223
 1224
 1225
 1226
 1227
 1228
 1229
 1230
 1231
 1232
 1233
 1234
 1235
 1236
 1237
 1238
 1239
 1240
 1241
 1242
 1243
 1244
 1245
 1246
 1247
 1248
 1249
 1250
 1251
 1252
 1253
 1254
 1255
 1256
 1257
 1258
 1259
 1260
 1261
 1262
 1263
 1264
 1265
 1266
 1267
 1268
 1269
 1270
 1271
 1272
 1273
 1274
 1275
 1276
 1277
 1278
 1279
 1280
 1281
 1282
 1283
 1284
 1285
 1286
 1287
 1288
 1289
 1290
 1291
 1292
 1293
 1294
 1295
 1296
 1297
 1298
 1299
 1300
 1301
 1302
 1303
 1304
 1305
 1306
 1307
 1308
 1309
 1310
 1311
 1312
 1313
 1314
 1315
 1316
 1317
 1318
 1319
 1320
 1321
 1322
 1323
 1324
 1325
 1326
 1327
 1328
 1329
 1330
 1331
 1332
 1333
 1334
 1335
 1336
 1337
 1338
 1339
 1340
 1341
 1342
 1343
 1344
 1345
 1346
 1347
 1348
 1349
 1350
 1351
 1352
 1353
 1354
 1355
 1356
 1357
 1358
 1359
 1360
 1361
 1362
 1363
 1364
 1365
 1366
 1367
 1368
 1369
 1370
 1371
 1372
 1373
 1374
 1375
 1376
 1377
 1378
 1379
 1380
 1381
 1382
 1383
 1384
 1385
 1386
 1387
 1388
 1389
 1390
 1391
 1392
 1393
 1394
 1395
 1396
 1397
 1398
 1399
 1400
 1401
 1402
 1403
 1404
 1405
 1406
 1407
 1408
 1409
 1410
 1411
 1412
 1413
 1414
 1415
 1416
 1417
 1418
 1419
 1420
 1421
 1422
 1423
 1424
 1425
 1426
 1427
 1428
 1429
 1430
 1431
 1432
 1433
 1434
 1435
 1436
 1437
 1438
 1439
 1440
 1441
 1442
 1443
 1444
 1445
 1446
 1447
 1448
 1449
 1450
 1451
 1452
 1453
 1454
 1455
 1456
 1457
 1458
 1459
 1460
 1461
 1462
 1463
 1464
 1465
 1466
 1467
 1468
 1469
 1470
 1471
 1472
 1473
 1474
 1475
 1476
 1477
 1478
 1479
 1480
 1481
 1482
 1483
 1484
 1485
 1486

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted August 21, 2003. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

us-09-761-466-6.rge

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Search time: 2463.3 seconds

Score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SEQUENCES

Result No.	Score	Query Match	Length	DB ID	Description
1	478	100.0	81512	9	A000412 Homo sapiens
2	466	97.5	14745	9	A000412 Homo sapiens
3	199.8	41.8	8117	10	A000412 Homo sapiens
4	199.8	41.8	15806	2	A000412 Homo sapiens
5	199.8	41.8	21918	2	A000412 Homo sapiens
6	182.2	48.1	8648	10	A000412 Homo sapiens
7	181.8	48.0	17110	2	A000412 Homo sapiens
8	181.8	48.0	241724	2	A000412 Homo sapiens
9	45.2	9.5	16317	2	A000412 Homo sapiens
10	45.2	9.5	19959	2	A000412 Homo sapiens
11	44.2	9.2	19144	10	A000412 Homo sapiens
12	42	8.8	14694	2	A000412 Homo sapiens
13	42	8.8	20715	2	A000412 Homo sapiens
14	41.8	8.7	7748	2	A000412 Homo sapiens
15	41.8	8.7	8543	9	A000412 Homo sapiens
16	41.8	8.7	11000	2	A000412 Homo sapiens
17	41.8	8.7	14174	9	A000412 Homo sapiens
18	41.8	8.7	22387	2	A000412 Homo sapiens
19	41.8	8.7	24826	2	A000412 Homo sapiens
20	41.8	8.7	28045	2	A000412 Homo sapiens
21	41.8	8.7	11000	2	A000412 Homo sapiens
22	41.2	8.6	24069	9	A000412 Homo sapiens
23	40.8	8.5	18464	10	A000412 Homo sapiens
24	40.8	8.5	18464	9	A000412 Homo sapiens
25	40.8	8.5	18464	9	A000412 Homo sapiens
26	40.8	8.5	19543	9	A000412 Homo sapiens
27	40.8	8.5	21974	10	A000412 Homo sapiens
28	40.4	8.5	6152	5	A000412 Homo sapiens
29	40.4	8.5	11000	2	A000412 Homo sapiens
30	40.4	8.5	15802	4	A000412 Homo sapiens
31	40.2	8.4	4848	2	A000412 Homo sapiens
32	40.2	8.4	12745	2	A000412 Homo sapiens
33	40.2	8.4	18217	2	A000412 Homo sapiens
34	40.2	8.4	20000	2	A000412 Homo sapiens
35	40.2	8.4	21244	2	A000412 Homo sapiens
36	40	8.4	18265	2	A000412 Homo sapiens
37	40	8.4	20019	10	A000412 Homo sapiens
38	40	8.4	20255	9	A000412 Homo sapiens
39	40	8.4	21756	2	A000412 Homo sapiens
40	39.8	8.3	24844	2	A000412 Homo sapiens
41	39.6	8.3	535	4	A000412 Homo sapiens
42	39.6	8.3	535	4	A000412 Homo sapiens
43	39.6	8.3	535	4	A000412 Homo sapiens
44	39.6	8.3	4770	9	A000412 Homo sapiens
45	39.6	8.3	14956	2	A000412 Homo sapiens

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	478	100.0	81512	9	A000412 Homo sapiens
2	466	97.5	14745	9	A000412 Homo sapiens
3	199.8	41.8	8117	10	A000412 Homo sapiens
4	199.8	41.8	15806	2	A000412 Homo sapiens
5	199.8	41.8	21918	2	A000412 Homo sapiens
6	182.2	48.1	8648	10	A000412 Homo sapiens
7	181.8	48.0	17110	2	A000412 Homo sapiens
8	181.8	48.0	241724	2	A000412 Homo sapiens
9	45.2	9.5	16317	2	A000412 Homo sapiens
10	45.2	9.5	19959	2	A000412 Homo sapiens
11	44.2	9.2	19144	10	A000412 Homo sapiens
12	42	8.8	14694	2	A000412 Homo sapiens
13	42	8.8	20715	2	A000412 Homo sapiens
14	41.8	8.7	7748	2	A000412 Homo sapiens
15	41.8	8.7	8543	9	A000412 Homo sapiens
16	41.8	8.7	11000	2	A000412 Homo sapiens
17	41.8	8.7	14174	9	A000412 Homo sapiens
18	41.8	8.7	22387	2	A000412 Homo sapiens
19	41.8	8.7	24826	2	A000412 Homo sapiens
20	41.8	8.7	28045	2	A000412 Homo sapiens
21	41.8	8.7	11000	2	A000412 Homo sapiens
22	41.2	8.6	24069	9	A000412 Homo sapiens
23	40.8	8.5	18464	10	A000412 Homo sapiens
24	40.8	8.5	18464	9	A000412 Homo sapiens
25	40.8	8.5	18464	9	A000412 Homo sapiens
26	40.8	8.5	19543	9	A000412 Homo sapiens
27	40.8	8.5	21974	10	A000412 Homo sapiens
28	40.4	8.5	6152	5	A000412 Homo sapiens
29	40.4	8.5	11000	2	A000412 Homo sapiens
30	40.4	8.5	15802	4	A000412 Homo sapiens
31	40.2	8.4	4848	2	A000412 Homo sapiens
32	40.2	8.4	12745	2	A000412 Homo sapiens
33	40.2	8.4	18217	2	A000412 Homo sapiens
34	40.2	8.4	20000	2	A000412 Homo sapiens
35	40.2	8.4	21244	2	A000412 Homo sapiens
36	40	8.4	18265	2	A000412 Homo sapiens
37	40	8.4	20019	10	A000412 Homo sapiens
38	40	8.4	20255	9	A000412 Homo sapiens
39	40	8.4	21756	2	A000412 Homo sapiens
40	39.8	8.3	24844	2	A000412 Homo sapiens
41	39.6	8.3	535	4	A000412 Homo sapiens
42	39.6	8.3	535	4	A000412 Homo sapiens
43	39.6	8.3	535	4	A000412 Homo sapiens
44	39.6	8.3	4770	9	A000412 Homo sapiens
45	39.6	8.3	14956	2	A000412 Homo sapiens

Prod. No. is the number of results produced by chance to have a

REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 81512)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
3 (bases 1 to 81512)
DOE Joint Genome Institute.
Direct Submission
Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 81512)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (01-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 81512)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (28-JAN-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.3% of Sequence;
Estimated Total Number of Errors is 0.4.
NOTE: This insert is not the entire sequence of the clone (entire sequence is 139.5kb). It is clipped at the overlaps with AC008378 and AC106731. The number of bases overlapped with AC008378 is 4442 bps and with AC106731 is 15413 bps.
Location/Qualifiers

FEATURES
source
1. .81512
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-281H14"

BASE COUNT 20011 a 19534 c 19980 g 21987 t
ORIGIN

Query Match 100.0%; Score 478; DB 9; Length 81512;
Best Local Similarity 100.0%; Pred. No. 3.8e-122; Indels 0; Gaps 0;
Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 AGAGAAATCATATACCGCATTCACAAGAGCATAGAGATGTAAACGATGATCTTGT 60
|||||
46088 AGAGAAATCATATACCGCATTCACAAGAGCATAGAGATGTAAACGATGATCTTGT 46029
|||||
61 CAAATAGGAGAGTTTTTTTCTTCCTTTTGTAAACACTGACCCACAGGACTGACAG 120
|||||
46028 CAAATAGGAGAGTTTTTTTCTTCCTTTTGTAAACACTGACCCACAGGACTGACAG 45969
|||||
121 TTCTAGGAGAGCCCTTACCGGAAATAGGAATTAATCTTGGCACCCTTGATTTGCAAG 180
|||||
45968 TTCTAGGAGAGCCCTTACCGGAAATAGGAATTAATCTTGGCACCCTTGATTTGCAAG 45909
|||||
181 GGCAATGCTAATTTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTTTTTAC 240
|||||
45908 GGCAATGCTAATTTTTTTCTTCTCCAGAGCTCTCAAAAAAATTTTTTTTAC 45849
|||||
241 TAAAAACAGGATCCCGATGATGCTGATGCTCCCAATTAACGATTAATTTTCAGGC 300
|||||
45848 TAAAAACAGGATCCCGATGATGCTGATGCTCCCAATTAACGATTAATTTTCAGGC 45789
|||||
301 GTCCGCTACACTAATCTTTTCAAAATGTCATCGCAGCCGCTGGCCACAGATTCACAT 360
|||||
45788 GTCCGCTACACTAATCTTTTCAAAATGTCATCGCAGCCGCTGGCCACAGATTCACAT 45729
|||||
361 AACAGCGCTCCAGAGACCTGCTCCGAGCTCTTTTCAGCGAGCAATTAATTTGAATCGG 420
|||||
45728 AACAGCGCTCCAGAGACCTGCTCCGAGCTCTTTTCAGCGAGCAATTAATTTGAATCGG 45669
|||||

OY 421 ATGTGCTCGTTTGGCAGAGCTACCGCTCGGCGATAGGATCTCTCCACGACAC 478
|||||
DB 45668 ATGTGCTCGTTTGGCAGAGCTACCGCTCGGCGATAGGATCTCTCCACGACAC 45611
|||||
RESULT 2
AL669924/C
LOCUS
DEFINITION
Human DNA sequence from clone XHbac-2L19 on chromosome 6, complete sequence.
ACCESSION
AL669924
VERSION
AL669924.9 GI:20068712
KEYWORDS
HTG.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
1
Peck, A.
Direct Submission
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19572797.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
XHbac-2L19 is from a CHO1-501 human bac - pgf cell line library VECTOR: pTARbac2.1
This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/chr6/MHC>.
Location/Qualifiers

FEATURES
source
1. .137345
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="XHbac-2L19"
/clone_1lb="CHO1-501"

BASE COUNT 33377 a 33207 c 33325 g 37436 t
ORIGIN

Query Match 97.5%; Score 466; DB 9; Length 137345;
Best Local Similarity 99.8%; Pred. No. 8.7e-119;
Matches 477; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
1 AGAGAAATCATATACCGATTCACAAGAGCATAGAGATGTAACATGATCTTGT 60
|||||
46314 AGAGAAATCATATACCGATTCACAAGAGCATAGAGATGTAACATGATCTTGT 46255
|||||
61 CAAATAGGAGAGTTTTTTTCTTCCTTTTGTAAACCTGACCCACAGAGCTACAG 120
|||||
46254 CAAATAGGAGAGTTTTTTTCTTCCTTTTGTAAACCTGACCCACAGAGCTACAG 46195
|||||
121 TTCTAGGAGAGCCCTTACCGCAAAATAGGAATAATCTTGGCCACTTGATTTGCAAG 180
|||||

14	46194	CGAAGAAATGCTGAGGAGAAATAGAAATAAATGCTGCTGCTGCTATTTGGAA	46195
15	181	GGCAAGATGATTTTCTTTCTTGACAGGCTTTAAAAAAGAAAAAAGCTTAC	210
16	46144	GGCAATGTAAATTTCTTTCTTGACAGGCTTG-AAAAAAGAAAAAAGCTTAC	45076
17	241	AAAAATATTAATTTTAACTAGGCTGATTTCTGATATAAGAGTAAATTTGAGT	30
18	46075	AAAAATATTAATTTTAACTAGGCTGATTTCTGATATAAGAGTAAATTTGAGT	45126
19	612	GTCATTTATATAATTTTAAATTTTATGTAAGAGGCTGGGAGAGATTTGACT	60
20	46015	GTCATTTATATAATTTTAAATTTTATGTAAGAGGCTGGGAGAGATTTGACT	45976
21	361	AACAGATGCTGATATAATGCTGAGGCTTTTATAGGACATTAATTTAAAGG	420
22	45965	AACAGATGCTGATATAATGCTGAGGCTTTTATAGGACATTAATTTAAAGG	45896
23	421	ATGAGGCTTTTATATAATGAGGCTGATAGAGATGCTGAAATATAT	478
24	45945	ATGAGGCTTTTATATAATGAGGCTGATAGAGATGCTGAAATATAT	45848

[illegible][illegible]

RESULT	4
A114621	
TITLE	MUS MUSCULUS CHROMOSOME DNA CONTENT ESTIMATION BY FLUORESCENT MICROANALYSIS
DEFINITION	
ABSTENTION	
VERSION	
KEYWORDS	HIS; HIS_PHASE; HIS_PAGE;
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathia; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 158809) McPherson J.D. and Watson R.H. The genome of Mus musculus clone unpublished 2 (bases 1 to 158809) McPherson J.D. and Watson R.H. Unpub Submission Submitted (6-MAY-2004) Genomic Sequencing Center 4444 University Ave Berkeley, CA 94720-5080 USA
COMMENT	
	Genomic Center Center: Washington University Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu/gen/index.shtml Contact: Submissions: wd.schwaner@wustl.edu Project: shihua@wustl.edu Center project name: Mmuschr04
	Summary Statistics Sequenced vector: 673120 Sequenced vector: assembled: 673120

Chemistry: Dye-Primer ET; 0% of reads
Chemistry: Dye-Terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 140428 bases at least Q40
Consensus quality: 146442 bases at least Q30
Consensus quality: 149661 bases at least Q20
Insert size: 217000; agarose-fp
Insert size: 153246; sum-of-contigs
Quality coverage: 1.93 in Q20 bases; agarose-fp
Quality coverage: 2.65 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 69 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1262: contig of 1262 bp in length
* 1263 1362: gap of unknown length
* 1363 2809: contig of 1447 bp in length
* 2810 2909: gap of unknown length
* 2910 4113: contig of 1204 bp in length
* 4114 4213: gap of unknown length
* 4214 5492: contig of 1279 bp in length
* 5493 5592: gap of unknown length
* 5593 6902: contig of 1310 bp in length
* 6903 7002: gap of unknown length
* 7003 8255: contig of 1253 bp in length
* 8256 8355: gap of unknown length
* 8356 9423: contig of 1068 bp in length
* 9424 9523: gap of unknown length
* 9524 10777: contig of 1254 bp in length
* 10778 10877: gap of unknown length
* 10878 12023: contig of 1146 bp in length
* 12024 12123: gap of unknown length
* 12124 13409: contig of 1286 bp in length
* 13410 13509: gap of unknown length
* 13510 15253: contig of 1744 bp in length
* 15254 15353: gap of unknown length
* 15354 16649: contig of 1296 bp in length
* 16650 16749: gap of unknown length
* 16750 18229: contig of 1480 bp in length
* 18230 18329: gap of unknown length
* 18330 19662: contig of 1333 bp in length
* 19663 19762: gap of unknown length
* 19763 21024: contig of 1262 bp in length
* 21025 21124: gap of unknown length
* 21125 22467: contig of 1343 bp in length
* 22468 22567: gap of unknown length
* 22569 23615: contig of 1048 bp in length
* 23616 23715: gap of unknown length
* 23716 25651: contig of 1936 bp in length
* 25652 25751: gap of unknown length
* 25752 27281: contig of 1530 bp in length
* 27282 27381: gap of unknown length
* 27382 29420: contig of 2039 bp in length
* 29421 29520: gap of unknown length
* 29521 30901: contig of 1381 bp in length
* 30902 31001: gap of unknown length
* 31002 32159: contig of 1158 bp in length
* 32160 32259: gap of unknown length
* 32260 33788: contig of 1529 bp in length
* 33789 33888: gap of unknown length
* 33889 35134: contig of 1246 bp in length
* 35135 35234: gap of unknown length
* 35235 36486: contig of 1252 bp in length
* 36487 36586: gap of unknown length
* 36587 38385: contig of 1799 bp in length
* 38386 38485: gap of unknown length
* 38486 39714: contig of 1229 bp in length
* 39715 39814: gap of unknown length

39815 41078: contig of 1264 bp in length
* 41079 41178: gap of unknown length
* 41179 42273: contig of 1095 bp in length
* 42274 42373: gap of unknown length
* 42374 44123: contig of 1750 bp in length
* 44124 44223: gap of unknown length
* 44224 45908: contig of 1685 bp in length
* 45909 46008: gap of unknown length
* 46009 47545: contig of 1537 bp in length
* 47546 47645: gap of unknown length
* 47646 50012: contig of 2367 bp in length
* 50013 50112: gap of unknown length
* 50113 51978: contig of 1866 bp in length
* 51979 52078: gap of unknown length
* 52079 54113: contig of 2035 bp in length
* 54114 54213: gap of unknown length
* 54214 55586: contig of 1373 bp in length
* 55587 55686: gap of unknown length
* 55687 57405: contig of 1719 bp in length
* 57406 57505: gap of unknown length
* 57506 58790: contig of 1285 bp in length
* 58791 58890: gap of unknown length
* 58891 60453: contig of 1563 bp in length
* 60454 60553: gap of unknown length
* 60554 62149: contig of 1536 bp in length
* 62150 62249: gap of unknown length
* 62250 63910: contig of 1661 bp in length
* 63911 64010: gap of unknown length
* 64011 65910: contig of 1900 bp in length
* 65911 66010: gap of unknown length
* 66011 67754: contig of 1744 bp in length
* 67755 67854: gap of unknown length
* 67855 69859: contig of 2005 bp in length
* 69860 71962: gap of unknown length
* 71963 72062: contig of 2003 bp in length
* 72063 74371: gap of unknown length
* 74372 74471: gap of unknown length
* 74472 76499: contig of 2028 bp in length
* 76500 76599: gap of unknown length
* 76600 78329: contig of 1730 bp in length
* 78330 78429: gap of unknown length
* 78430 81033: contig of 2604 bp in length
* 81034 81133: gap of unknown length
* 81134 82450: contig of 1317 bp in length
* 82451 82550: gap of unknown length
* 82551 85880: contig of 3330 bp in length
* 85881 85980: gap of unknown length
* 85981 89359: contig of 3979 bp in length
* 89360 90059: gap of unknown length
* 90060 92499: contig of 2440 bp in length
* 92500 92599: gap of unknown length
* 92600 95372: contig of 2773 bp in length
* 95373 95472: gap of unknown length
* 95473 98796: contig of 3324 bp in length
* 98797 98896: gap of unknown length
* 98897 101619: contig of 2733 bp in length
* 101620 101719: gap of unknown length
* 101720 105463: contig of 3744 bp in length
* 105464 105563: gap of unknown length
* 105564 108520: contig of 2957 bp in length
* 108521 108620: gap of unknown length
* 108621 111714: contig of 3094 bp in length
* 111715 111814: gap of unknown length
* 111815 115701: contig of 3887 bp in length
* 115702 115801: gap of unknown length
* 115802 117720: contig of 1919 bp in length
* 117721 117820: gap of unknown length
* 117821 121279: contig of 3459 bp in length
* 121280 121379: gap of unknown length
* 121380 125489: contig of 4110 bp in length
* 125490 125589: gap of unknown length
* 125590 128887: contig of 3308 bp in length

Db 20912 ATCCGAGGGCCAGCCGATTTCATTAAACAGACTCCAGAGACCTCCGCCAGAGCTGTTT 20971
QY 396 TCAGCGAGACATTTAATTGATCGATCGCTGCTTGGCCAGACGTCACCGCTCGGCG 455
Db 20972 TCACAGACGATTTAATTGATCGATCGCTGCTTGGCCAGACGTCACCGCTCGGCG 21031
QY 456 ATAGCATCTCTCCACGACAC 478
Db 21032 ATAGCATCTCTCCACGACAC 21054

RESULT 6
AF083133 8648 bp DNA linear ROD 08-DEC-2000
LOCUS Mus musculus cardiac homeobox transcription factor (Kkx2-5) gene,
DEFINITION complete cds.
ACCESSION AF083133
VERSION AF083133.1 GI:11602837
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Tanaka,M., Wechsler,S.B., Lee,I.W., Yamasaki,N., Lawlits,J. and
Izumo,S.
TITLE Modular cis-acting elements regulate expression of the murine
Csk/Kkx2-5 homeobox gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8648)
AUTHORS Lee,I.W., Wechsler,S.B., Tanaka,M., Yamasaki,N. and Izumo,S.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1998) Cardiology, Beth Israel Deaconess Medical
Center, Harvard Medical School, 330 Brookline Avenue, Boston, MA
02215, USA

FEATURES
source 1..8648 Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129"
/db_xref="taxon:10090"
1114..8114
/gene="Kkx2-5"
/note="synonym: Csk"
join(1114..1159,1885..1920,5171..5754,7132..8114)
/gene="Kkx2-5"
/product="cardiac homeobox transcription factor"
/note="alternatively spliced"
join(1828..1920,5171..5754,7132..8114)
/gene="Kkx2-5"
/product="cardiac homeobox transcription factor"
/note="alternatively spliced"
join(5100..5754,7132..8114)
/gene="Kkx2-5"
/product="cardiac homeobox transcription factor"
/note="alternatively spliced"
join(5424..5754,7132..7757)
/gene="Kkx2-5"
/function="involved in heart development"
/note="timman homolog; Kkx family"
/codon_start=1
/product="cardiac homeobox transcription factor"
/protein_id="AAG38875.1"
/db_xref="GI:11602838"
/translation="MPPSPALTPPFSYKDLINLEOQRLASGDLASRLAETLAPAS
CMLAFKPEASGEEMASGLAEELRAEMGRAPSPKCSPPAPATFPFGATGDDPA
KDPRAKKEELALQKAVELDAETDGARPPARRRRKRVLESOAVLELEKQQR
YLSAERQOLASVLTISTQYKIMFQNRKRCRQRQDTLELGGPPPPARLAVP
LVRGKPCIDGPAAYAPAYAGVGLNAYGYNAYPSYSGAACSPPGSCAAYAAPPAAH
APASANSFNPFVGDINTVQSPMPQNGSVSLHGIRAM"

BASE COUNT 2041 a 2336 c 2179 g 2091 t 1 others

Query Match 38.1%; Score 182.2; DB 10; Length 8648;
Best Local Similarity 71.6%; Pred. No. 9,6e-40;
Matches 317; Conservative 0; Mismatches 103; Indels 23; Gaps 5;

QY 36 GAGCTAACAGTACTGATCTTTGTCAAATAGGAGAGTTTTCCTTCCTTTTGT 95
Db 3009 GGGTGGGAAGACTGATGTTTGTCAATATGAAAGAGTTTTT-----CTTCCTTTTCTT 3063
QY 96 AACACCTGACCCAGAGAGTACAGTTCAGAGAGCCGCCCTTACGAAATAGGAATA 155
Db 3064 GACACCTGACCCAGAGAGTTCGTAACCTCT-6GAAGCCCTTATATCGAAAA-AAATGATA 3121
QY 156 AATCCTGCACTTATTTGCAAGGCAATCTAATTTTCTTCTTCAGAGCTCTC 215
Db 3122 AATCCTCAACATGATTTGCAAGGAAG-----TCCCTCTCAAACTAT 3171
QY 216 AAAAAAAAAAAAAAAAACTTACTAAAAACAGGATCCCGATGTAGCTTCGATGCC 275
Db 3172 TTTTAAAAAGACTTAAAAACAGGATCCCGATGGGGGCCCAATATGCTC-----CC 3225
QY 276 CCATTTAAAGGTATATTTTCAGGGTCCGCTCACACTAATCTTCAAACTGCATGCCG 335
Db 3226 CCATTTAAAGGTATATTTTCAGGGTCCGCTCACACTAATCTTCAAACTGCATGCCG 3285
QY 336 AGCCGCTGGCCAGCAGATTCCTTACAGCGCTCCAGAGACCTGCTTCGAGCTCTTT 395
Db 3286 ATCCGAGGGGAGCGAGATTCCTTACAGAGACTCCAGAGACCTGCGGAGCTGTTT 3345
QY 396 TCAGCGAGACATTTAATTGAATCGATGTGCTGCTTTCAGAGAGTCCGCCCTCGCGG 455
Db 3346 TCAGCGAGACATTTAATTGAATCGATGTGCTGCTTTCAGAGAGTCCGCCCTCGCGG 3405

QY 456 ATAGCATCTCTCCACGACAC 478
Db 3406 ATAGCATCTCTCCACGACAC 3428

RESULT 7
AC119699/AC119699
LOCUS AC119699.8 GI:24941740
DEFINITION Rattus norvegicus clone CH230-466D21, WORKING DRAFT SEQUENCE, 5
unoriented pieces.
ACCESSION AC119699
VERSION AC119699.8 GI:24941740
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 171110)
Muzny,D.,Marie., Metzker,M.,Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebech,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Bisvalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Gantla,R., Garcia,A., Garner,T., Garza,M.,
Gebrageorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haylak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Huylk,S., Hume,J., Idlebird,R., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,D., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,

AC102990/c	LOCUS	AC102990	241724 bp	DNA	linear	HTG 10-MAY-2003	
	DEFINITION	Rattus norvegicus clone CH230-53024, WORKING DRAFT SEQUENCE.					
	ACCESSION	AC102990					
	VERSION	AC102990.6	GI:30521867				
	KEYWORDS	HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.					
	SOURCE	Rattus norvegicus (Norway rat)					
	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
REFERENCE	AUTHORS	1 (bases 1 to 241724) Muzny,D,Marle, Metzker,M,lee, Abramzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amlin,A, Angiaron,D, Anyalbebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Beha,B, Bistwalto,K, Blair,T, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buha,Y, C, Burch,P, Butrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Ceaar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,T, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Ande,C, Dedetich,D, Delgado,O, Denison,S, Detamo,C, Ding,X, Dim,H, Dlyva,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durkin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Georegeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamli,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Huluj,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Kapaphy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kwis,C, Krieff,C,L, Lebow,H, Levan,T, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenschewa,L, Louised,H, Lozaro,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Manghney,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwokoleh,O, Okwunu,G, Olarpunsgoon,A, Pal,S, Parks,K, Patelnek,S, Paul,H, Perez,A, Perez,L, Pfankech,C, Plopper,F, Polndexter,A, Popovic,D, Prims,E, Pu,L, L, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Rejler,M, Reigh,R, Rellly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smaj,D, Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,M, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanat,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wlecyzk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakud,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhauser,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.					
TITLE	JOURNAL	Unpublished					
REFERENCE	AUTHORS	2 (bases 1 to 241724) Worley,K.C.					
TITLE	JOURNAL	Direct Submission					
REFERENCE	AUTHORS	Submitted (24-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA					
TITLE	JOURNAL	3 (bases 1 to 241724) Rat Genome Sequencing Consortium.					
REFERENCE	AUTHORS	Submitted (10-MAY-2003) Human Genome Sequencing Center, Department					

COMMENT	
	of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
	On May 10, 2003 this sequence version replaced gi:22855562.
	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
	----- Genome Center -----
	Center: Baylor College of Medicine
	Center code: BCM
	Web site: http://www.hgsc.bcm.tmc.edu/
	Contact: hgsc-help@bcm.tmc.edu
	----- Project Information -----
	Center project name: GHEJ
	Center clone name: CH230-53024
	----- Summary Statistics -----
	Assembly program: Atlas 3.0;
	Consensus quality: 233518 bases at least Q40
	Consensus quality: 235377 bases at least Q30
	Consensus quality: 236878 bases at least Q20
	Estimated insert size: 245960; sum-of-contigs estimation
	Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE:	Estimated insert size may differ from sequence length
*	(see http://www.hgsc.bcm.tmc.edu/docs/genbank-draft_data.html).
* NOTE:	This is a 'working draft' sequence.
*	consists of 1 contigs. Gaps between the contigs
*	are represented as runs of N. The order of the pieces
*	is believed to be correct as given, however the sizes
*	of the gaps between them are based on estimates that have
*	provided by the submitor.
*	This sequence will be replaced
*	by the finished sequence as soon as it is available and
*	the accession number will be preserved.
1	241724: contig of 241724 bp in length.
FEATURES	
source	location/Qualifiers
	1..241724
	/organism="Rattus norvegicus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10116"
	/clone="CH230-53024"
	1..2211
misc_feature	/note="wgs_contig"
BASE COUNT	60533 a 53609 c 56345 g 67996 t 3241 others
ORIGIN	
Query Match	38.0%; Score 181.8; DB 2; Length 241724;
Best Local Similarity	68.6%; Pred. No. 1.5e-39;
Matches 319;	Conservative 0; Mismatches 122; Indels 24; Gaps 4;
OY	15 CCGATTCAACAAGAGCATGAGAGTGTAACTACTGATCTTTGTCACAAATAGGGAGACT 74
DB	168674 CCTTTTACACCAAGTGTTTAGAGGGCGGGAAGCTCATACTGATTTGTCCAATTAAGAAGATT 168615
OY	75 TTTTTTTCCTCCCTTTTGTAAACCCTGACCCACAGAGCTGCACATTCCTAGGAACGCC 134
DB	168614 TTT-----CTTCTTTTTCCTTACACCTGACCCACAGATCTGTCAACTCT-GGAGACCTTT 168561
OY	135 CTTACCCGAAATATAGCAATTAATCTCTTGCACCTTGATTTGCACAGGCATCTAATT 194
DB	168560 TATATCCGGGAAAAGTGATTAATCCCCCACAFATGATTGGAAAGGAATGTCCCTC 168501
OY	135 TTTTCTTTCTCCAGAGCTCTCAAAAAAAAAAAAAAACCTTACTAAAAACAGSGATC 254
DB	168500 TCAGACATCATCTTTTAAAAAAGAGCTTAAACAGGAGATCCCGATGTGGCCC 168441

[illegible]

```

COMMENT
TITLE
JOURNAL
DIRECT SUBMISSION
Submitted (08-03-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 6, 2002 this sequence version replaced g2:21700716.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/MW/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WBX
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 125795
Center clone name: 531_C8
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: dideoxynucleotide Big Dye; 100% of reads
Assembly program: Phrap; version 0.999741
Consensus quality: 156910 bases at least 40
Consensus quality: 161126 bases at least 92
Insert size: 17000-42000 bp
Insert size: 161547; sum-of-reads
quality coverage: 4.7 in 920 bases; adarose-1p
quality coverage: 5.0 in 920 bases; sum-of-reads
-----
NOTE: This is a "working draft" sequence. It currently
* consists of 16 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1
1213: contig of 1212 bp in length
1313 1212: contig of 1212 bp in length
1313 2590: contig of 1478 bp in length.
2991 3090: gap of 100 bp
3091 5404: contig of 2314 bp in length
5405 5505: gap of 100 bp
5505 8570: contig of 3065 bp in length
8571 8671 12334: contig of 3664 bp in length.
12345 12434: gap of 100 bp
12435 17442: contig of 4908 bp in length
17443 21556: contig of 4214 bp in length.
21557 21556: gap of 100 bp
21557 25915: contig of 4459 bp in length
25916 26015: gap of 100 bp
26016 32932: contig of 6917 bp in length
32933 33032: gap of 100 bp
33033 40214: contig of 7182 bp in length
40215 40314: gap of 100 bp
40315 50386: contig of 10072 bp in length.
50387 50486: gap of 100 bp
50487 65419: contig of 14884 bp in length
65420 65420 79408: contig of 13989 bp in length
79409 79508: gap of 100 bp
79509 98119: contig of 16611 bp in length
98120 98219: gap of 100 bp
98220 113478: contig of 17759 bp in length
113479 113478: gap of 100 bp
113479 161379: contig of 48079 bp in length.

```


Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
-----Project Information-----
Center project name: L6875
Center clone name: 404_F_23

* NOTE: This record contains 85 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1 783: contig of 783 bp in length
* 784 883: gap of 100 bp
* 884 1715: contig of 832 bp in length
* 1716 1815: gap of 100 bp
* 1816 2635: contig of 820 bp in length
* 2636 2735: gap of 100 bp
* 2736 3523: contig of 788 bp in length
* 3524 3623: gap of 100 bp
* 3624 4437: contig of 814 bp in length
* 4438 4537: gap of 100 bp
* 4538 5363: contig of 826 bp in length
* 5364 5463: gap of 100 bp
* 5464 6265: contig of 802 bp in length
* 6266 6365: gap of 100 bp
* 6366 7189: contig of 824 bp in length
* 7190 7289: gap of 100 bp
* 7290 8104: contig of 815 bp in length
* 8105 8204: gap of 100 bp
* 8205 9008: contig of 804 bp in length
* 9009 9108: gap of 100 bp
* 9109 9939: contig of 831 bp in length
* 9940 10039: gap of 100 bp
* 10040 10873: contig of 834 bp in length
* 10874 10973: gap of 100 bp
* 10974 11801: contig of 828 bp in length
* 11802 11901: gap of 100 bp
* 11902 12697: contig of 796 bp in length
* 12698 12797: gap of 100 bp
* 12798 13591: contig of 794 bp in length
* 13592 13691: gap of 100 bp
* 13692 14505: contig of 814 bp in length
* 14506 14605: gap of 100 bp
* 14606 15432: contig of 827 bp in length
* 15433 15532: gap of 100 bp
* 15533 16360: contig of 828 bp in length
* 16361 16460: gap of 100 bp
* 16461 17278: contig of 818 bp in length
* 17279 17378: gap of 100 bp
* 17379 18214: contig of 836 bp in length
* 18215 18314: gap of 100 bp
* 18315 19152: contig of 838 bp in length
* 19153 19252: gap of 100 bp
* 19253 20088: contig of 836 bp in length
* 20089 20188: gap of 100 bp
* 20189 20963: contig of 775 bp in length
* 20964 21063: gap of 100 bp
* 21064 21894: contig of 831 bp in length
* 21895 21994: gap of 100 bp
* 21995 22759: contig of 765 bp in length
* 22760 22859: gap of 100 bp
* 22860 23645: contig of 786 bp in length
* 23646 23745: gap of 100 bp
* 23746 24541: contig of 796 bp in length
* 24542 24641: gap of 100 bp

24642 25447: contig of 806 bp in length
* 25448 25547: gap of 100 bp
* 25548 26373: contig of 826 bp in length
* 26374 26473: gap of 100 bp
* 26474 27287: contig of 814 bp in length
* 27288 27387: gap of 100 bp
* 27388 28210: contig of 823 bp in length
* 28211 28311: gap of 100 bp
* 28311 29107: contig of 797 bp in length
* 29108 29207: gap of 100 bp
* 29208 30038: contig of 831 bp in length
* 30039 30138: gap of 100 bp
* 30139 30966: contig of 828 bp in length
* 30967 31066: gap of 100 bp
* 31067 31872: contig of 806 bp in length
* 31873 31972: gap of 100 bp
* 31973 32772: contig of 800 bp in length
* 32773 32872: gap of 100 bp
* 32873 33672: contig of 800 bp in length
* 33673 33772: gap of 100 bp
* 33773 34587: contig of 815 bp in length
* 34588 34687: gap of 100 bp
* 34688 35506: contig of 819 bp in length
* 35507 35606: gap of 100 bp
* 35607 36418: contig of 812 bp in length
* 36419 36518: gap of 100 bp
* 36519 37340: contig of 822 bp in length
* 37341 37440: gap of 100 bp
* 37441 38266: contig of 826 bp in length
* 38267 38366: gap of 100 bp
* 38367 39191: contig of 825 bp in length
* 39192 39291: gap of 100 bp
* 39292 40075: contig of 784 bp in length
* 40076 40175: gap of 100 bp
* 40176 41005: contig of 830 bp in length
* 41006 41105: gap of 100 bp
* 41106 41919: contig of 814 bp in length
* 41920 42822: contig of 803 bp in length
* 42823 42922: gap of 100 bp
* 42923 43745: contig of 823 bp in length
* 43746 43845: gap of 100 bp
* 43846 44653: contig of 808 bp in length
* 44654 44753: gap of 100 bp
* 44754 45563: contig of 810 bp in length
* 45564 45663: gap of 100 bp
* 45664 46491: contig of 828 bp in length
* 46492 46591: gap of 100 bp
* 46592 47432: contig of 841 bp in length
* 47433 47532: gap of 100 bp
* 47533 48374: contig of 842 bp in length
* 48375 48474: gap of 100 bp
* 48475 49273: contig of 799 bp in length
* 49274 49373: gap of 100 bp
* 49374 50190: contig of 817 bp in length
* 50191 50290: gap of 100 bp
* 50291 51083: contig of 793 bp in length
* 51084 51183: gap of 100 bp
* 51184 51939: contig of 756 bp in length
* 51940 52039: gap of 100 bp
* 52040 52848: contig of 809 bp in length
* 52849 52948: gap of 100 bp
* 52949 53763: contig of 815 bp in length
* 53764 53863: gap of 100 bp
* 53864 54698: contig of 835 bp in length
* 54699 54798: gap of 100 bp
* 54799 55573: contig of 775 bp in length
* 55574 55673: gap of 100 bp
* 55674 56508: contig of 835 bp in length
* 56509 57439: gap of 100 bp
* 57440 57539: contig of 831 bp in length
* 57540 58355: contig of 816 bp in length

[illegible][illegible]

```
repeat_region      5587. .6005  
/rpt_family="ERV1"  
repeat_region      6006. .6518  
/rpt_family="MALR"  
repeat_region      6519. .7199  
/rpt_family="ERV1"  
repeat_region      7200. .7494  
/rpt_family="Alu"  
repeat_region      7495. .7987  
/rpt_family="ERV1"  
repeat_region      8014. .8507  
/rpt_family="ERV1"  
repeat_region      8557. .8948  
/rpt_family="MALR"  
repeat_region      9002. .9396  
/rpt_family="ERV1"  
repeat_region      10045. .10493  
/rpt_family="ERV1"  
repeat_region      11114. .11591  
/rpt_family="MALR"  
repeat_region      12897. .13024  
/rpt_family="L1"  
repeat_region      13078. .13108  
/rpt_family="AT_rich"  
repeat_region      13218. .13321  
/rpt_family="MALR"  
repeat_region      13332. .13662  
/rpt_family="MALR"  
repeat_region      15233. .15565  
/rpt_family="MALR"  
repeat_region      15366. .15739  
/rpt_family="MALR"  
repeat_region      16034. .16063  
/rpt_family="AT_rich"  
repeat_region      16114. .16158  
/rpt_family="(TG)n"  
repeat_region      16365. .16827  
/rpt_family="ERV1"  
repeat_region      16885. .16922  
/rpt_family="(CA)n"  
repeat_region      21089. .21516  
/rpt_family="ERV1"  
repeat_region      21517. .21812  
/rpt_family="Alu"  
repeat_region      21813. .22058  
/rpt_family="ERV1"  
repeat_region      22059. .22424  
/rpt_family="L1"  
repeat_region      22425. .23556  
/rpt_family="MER2_type"  
repeat_region      23557. .24983  
/rpt_family="L1"  
repeat_region      25147. .25565  
/rpt_family="L1"  
repeat_region      25584. .26360  
/rpt_family="L1"  
repeat_region      26367. .27217  
/rpt_family="L1"  
repeat_region      27255. .27587  
/rpt_family="L1"  
repeat_region      27588. .27866  
/rpt_family="Alu"  
repeat_region      27870. .27891  
/rpt_family="(TAAAA)n"  
repeat_region      27892. .28015  
/rpt_family="L1"  
repeat_region      28758. .29098  
/rpt_family="Alu"  
repeat_region      30348. .30656  
/rpt_family="Alu"  
repeat_region      30779. .30799  
/rpt_family="AT_rich"  
repeat_region      31174. .31384
```

```
repeat_region      /rpt_family="Alu"  
31697. .31896  
/rpt_family="CRL"  
repeat_region      31904. .31951  
/rpt_family="ERV1"  
repeat_region      32005. .32473  
/rpt_family="L1"  
repeat_region      32598. .32707  
/rpt_family="L1"  
repeat_region      32792. .32874  
/rpt_family="(TA)n"  
repeat_region      32992. .33026  
/rpt_family="AT_rich"  
repeat_region      33571. .33598  
/rpt_family="AT_rich"  
repeat_region      33665. .33708
```

Query Match 8.7%; Score 41.8; DB 9; Length 85439;
Best Local Similarity 51.3%; Pred. No. 1.4;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```
QY 57 TGTCAATAAGGAGAGTTTTCCTTCCTTTTGTAAACCTGACCCACAGGACG 116  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1454 TTTCTAAACAAAACACATTTTGCTGCTCTATTAACATGTCATCACAACAGGTCCTG 1513  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 117 ACAATTCTAGGAAGCCCTTACCCGAAATAGAAATAAATCCTTGCCACCTGATTTG 176  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1514 ACAGGACATATATCCACCTCTCTATATATAGTGAAAAAAATATTAAGACTCTAAATTGAA 1573  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 177 CAAGGCAATGCTAATTTTCTTCCTTCCTCCAGAGCTCCGCAAAAAA 236  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1574 AGAGGGAAACTAATACATATGTTATATACAAAAGCTCTCCACATATTTAAAGTGAAC 1633  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
QY 237 TTAATAAAA 245  
| | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1634 TTAGTCAAA 1642
```

Search completed: August 20, 2003, 12:34:45
Job time : 2469.3 secs


```
source
1. .577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HFWAHE01"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="HTF"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      165 a      152 c      157 g      103 t
ORIGIN

Query Match      24.1%; Score 258; DB 9; Length 577;
Best Local Similarity 87.6%; Pred. No. 6.7e+02;
Matches 282; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 626 GGCCGGGGCCATGCTCAGCCTGTATATCCAGCACTTTAGAGGCGGAGCGGCCGA 685
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 70 GGCCGGGGCGGTGGCTCAGCCTGTATATCCAGCACTTTGGAGGCGGAGCGGCCAGA 129
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 686 TCACCTGAGTGGGAGTTGCGACACCGCTGGCCACATGATGTGAACCTGTCTCTACA 745
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 130 TCACCTGAGTGGGAGTTGCGACACCGCTGGCCACATGATGTGAACCTGTCTCTACT 189
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 746 AAAAATACAAAATTAGCGCGGCGATGGTATGCTGATCCGCTGATCCAGTACTCGGAG 805
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 190 AAAAATACAAAATTAGCGCGGCGATGGTATGCTGATCCGCTGATCCAGTACTCGGAG 249
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 806 GCTGAGGAGAGATGCTTGAACCGGAGCGGAGGTTGCACTGAGCCGAGATCACA 865
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 250 GCGGAGGAGAGATGCTTGAACCGGAGCGGAGGTTGCACTGAGCCGAGATCAGG 309
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 866 CCAGTGCACCTCCAGCTGGGCGACAAAGCGAAATTCGCTCAAAAAATTAATAAAT 925
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 310 CCAGTGCACCTCCAGCTGGGCGACAAAGCGAAATTCGCTCAAAAAATTAATAAAT 369
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 926 AAAATGATATTAAGCCCATCA 947
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 370 AGAAAAAGAAATTAGCCAGGCA 391
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 2
AV730440      661 bp      mRNA      linear      EST 17-OCT-2000
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 661)
Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu
S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
Chen,J., Chen,Z. and Han,Z.
Homo sapiens cDNA HTF clones
Unpublished
JOURNAL
COMMENT
Contact: Zenguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1. .661
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="HFWAHE01"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="HTF"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      183 a      178 c      179 g      116 t      5 others
ORIGIN

Query Match      24.1%; Score 258; DB 9; Length 661;
Best Local Similarity 87.6%; Pred. No. 6e+02;
Matches 282; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 626 GGCCGGGGCCATGCTCAGCCTGTATATCCAGCACTTTAGAGGCGGAGCGGCCGA 685
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 70 GGCCGGGGCGGTGGCTCAGCCTGTATATCCAGCACTTTGGAGGCGGAGCGGCCAGA 129
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 686 TCACCTGAGTGGGAGTTGCGACACCGCTGGCCACATGATGTGAACCTGTCTCTACA 745
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 130 TCACCTGAGTGGGAGTTGCGACACCGCTGGCCACATGATGTGAACCTGTCTCTACT 189
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 746 AAAAATACAAAATTAGCGCGGCGATGGTATGCTGATCCGCTGATCCAGTACTCGGAG 805
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 190 AAAAATACAAAATTAGCGCGGCGATGGTATGCTGATCCGCTGATCCAGTACTCGGAG 249
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 806 GCTGAGGAGAGATGCTTGAACCGGAGCGGAGGTTGCACTGAGCCGAGATCACA 865
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 250 GCGGAGGAGAGATGCTTGAACCGGAGCGGAGGTTGCACTGAGCCGAGATCAGG 309
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 866 CCAGTGCACCTCCAGCTGGGCGACAAAGCGAAATTCGCTCAAAAAATTAATAAAT 925
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 310 CCAGTGCACCTCCAGCTGGGCGACAAAGCGAAATTCGCTCAAAAAATTAATAAAT 369
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY 926 AAAATGATATTAAGCCCATCA 947
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 370 AGAAAAAGAAATTAGCCAGGCA 391
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

RESULT 3
AM469140/c      437 bp      mRNA      linear      EST 24-FEB-2000
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 437)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 420.

FEATURES
source
1. .437
Location/Qualifiers
/organism="Homo sapiens"
```


Oy	714	CGTGGCCCAACATGCTGTAAGAACCCCTGCTCTACAAAAAAATACAAAAATTAAGCCGGCATGCT	773
Db	275	CGTGGCCCAACATGCTGTAAGAACCCCTGCTCTACTAAAAAACAAAAATTAAGCCAGGCGGGCT	216
Oy	774	GATCGGTCCTGTGATCCCACTACTCGGAGGCTGAGGACGAGAATTCGTTGAACCCG	833
Db	215	GGTGGTGTCCTGTATATCCCACTTCTTGAGGCTGAGGCGAGGAATTCACCTTGAACCCG	156
Oy	834	GGAGGCGGAGGTTGCAGAGGACGGAGATACACACATGCACCTCCAGCCTGGGCGCACAAGA	893
Db	155	GGAGGCGGAGGTTGCAGTACGTACCTAGATCGCGCCACACGCACTCCAGCCTGGGCACTACA	96
Oy	894	GCGAAATTCGCTCTAAAAAAATTAATTAATTAATA	928
Db	95	CCGAAACTCGCTCTCAAAAAAAGAAAAAAGAAAA	61
RESULT 9	AL708218	690 bp	linear
LOCUS	AL708218	690 bp	linear
DEFINITION	DKFZp686M2350.11 686 (synonym: h1c3) Homo sapiens cDNA clone		
ACCESSION	AL708218		
VERSION	AL708218.1	GI:19691573	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Ottewaelder, B., Obermaier, B., Mewes, W., Mewes, H.W., Well, B. and Wiemann, S.		
TITLE	EST (Ottewaelder, B., Obermaier, B., Mewes, H.W., Well, B. and Wiemann, S.)		
JOURNAL	Unpublished		
COMMENT	Contact: Ottewaelder B		
	MIPS		
	IngoIsaetler Landstr.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		
	sequenced by MedGenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available.		
	This clone (DKFZp686M2350) is available at the RZPD in Berlin.		
	Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059 Berlin-Charlottenburg; KERNMAY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..690		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DKFZp686M2350"		
	/tissue_type="human skeletal muscle"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_lib="686 (synonym: h1c3)"		
	/note="vector: pRipilex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"		
BASE COUNT	162 a 185 c 225 g 118 t		
ORIGIN			
Query Match	23.2%	Score 248.6;	DB 9;
Best Local Similarity	86.2%;	Pred. No. 8.8e+02;	
Matches	275;	Conservative 0;	Mismatches 44;
		Indels 0;	Gaps 0;
Oy	605	TAGCGAATTTTAAGACACATCAGCGCGGCGGCATGCTCAGCGCTGTATCCAGCACTT	664
Db	28	TAGAGCCTTTAAAGACGACCTGCGCGCGCGGTGCTCAGCGCTGTATCGACACACTT	87
Oy	665	TAGAGCGCGGAGCGGCGGCATCAGTCTAGTCTGGGAATTTGACACCAAGCTCGGCACAA	724
Db	88	GGGAGGCGCGGAGCGGCGGATCATCTTAGGTCAACAATTTGAGAGCCAGCTGGGCCAAC	147

OY		725 TGGGAACACCCTGTCCTACTACAAAATAATTAAGCCGGCGCATGGTATGCCTGT	784
Db		148 CGCGAACAACCTCCTCTCTACTAAAAATNCAAAATTATAGCTTGATGGTGGTCTCT	207
OY		785 GTGATCCCAGCTACTCTCGGAGGCTGTAGGACAGAGAATTCGTTGAACCCGGAGGCCAGG	844
Db		208 GTAATCCCACTACTCAGAGGCTGTAGACAGAGAAATCGCTTGAACTGGGAGGAGGAGG	267
OY		845 TTTCAGTAGCACCGAGATCAACACCACTCCAGCTGGGGCGCAAGAGCAAATTCGCG	904
Db		268 TTGGAGTAGAACCGAGATGTGTACCTTCCTCACTCCAGCTCGGCGCAAGAGCAAATTCGCG	327
OY		905 TCATAAAAAAAAAATAATATA 923 	
Db		328 TTCATAAGAAAAGAAAGANA 346	
RESULT 10			
BMA60657 LOCUS DEFINITION		BMA60657 1136 bp mRNA linear EST 05-FEB-2002 AGNCOURT 6421229 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532003 5'. mRNA sequence.	
ACCESSION VERSION KEYWORDS SOURCE ORGANISM		BMA60657 BMA60657.1 GI:18509697 EST. Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Catarrhini; Homnidae; Homo.	
REFERENCE AUTHORS TITLE JOURNAL COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC) National Institute of Health, Ph.D. Contact: Robert Strausberg, Ph.D. Email: cgabs@email.nih.gov Tissue procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLB) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov Plate: LLAMJ2214 row: j column: 04 High quality sequence stop: 534. Location/Qualifiers 1..1136 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5532003" /_tisue_type="_telomysatcoma" /_lab_host="DH10B (phage-resistant)" /_note="Organ: uterus; Vector: pCMV-Sport6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: oligo dT. Average insert size 2.1 Kb." Average insert size 2.1 Kb."	
FEATURES source			
Query Match Best local Similarity Matches		23.2%; Score 248.6; DB 12; Length 1136; 86.2%; Pred. No. 6e+02; Conservative 0;	Mismatches 44; Indels 0; Gaps 0;
OY Db OY Db OY Db OY		605 TAGCGAATTTAAACACATCATGCGCGCGCGCGCATAGCGTCACGCTGTAAATCCAGCACTT 134 TAGAGCCTTTGAAAACGAGCCTGCGCGCGCGCGGTGCGTCAAGCCTGTATCGCAGCACTT 665 TAGAGGCGCGAGCGCGCGCGCATCACTCTGAGGTTCGGAGTTCGACACAGCCTGGCCAAMA Db 194 GGSGAGGCGGAGCGAGCGGAGATCACTGAGGTGAGAAAGTTGCGAGCCAGCGCTGGCCAAMA 253 OY 725 TGCTGAACCCCTGYCTCTACAAAAAATNCAAAATTTAGCCGGCGCATGGTATGCCTGTCT	664 193 724 253 784

[illegible][illegible]

Qy	669	GGGGCGAGCGGGGGCGGATCCACTGAGGTCGGAGTTGCAACACAGCTGGCCACATAGT	728
Db	596	AGCGCGAGGTGGGTGATCACCCTGAGGTGAGGATTTGAGACCAACCTGGCCAAATAGT	655
Qy	729	GAACCCCTGTCTACAAATAATACAAAAATTAGCCGGGCGATGATGCGTCCCTGTGA	788
Db	656	GAACCCCGTCTCTACTAAAAATACAAAAGTTAGCCGGGCAATGAGTGTGCGCTGTAA	715
Qy	789	TCCCACTACTCGGAGAGCTGAGGAGAGAAATCGCTTGAACCCGGGAGGGGAGTGGC	848
Db	716	TCCCACTACTCAGGAGGCTTANAGGAGGAATCGCTTGAACCCGAAAGGCGGAATTC	775
Qy	849	AGTAGCCGAGATCACACCACTGCCTCCAGCTGGGGACAAAGACGAAATTCCTTA	908
Db	776	AGTAGCCAAAGATCGTGGCCTGCACCTGCACCTCGAGCTGGGCAACAAGACAAAGACTGTCTC	835
Qy	909	AAAAAATAAATAAATAAATGA	932
Db	836	AAAAAATAATAAATAAATTA	859

RESULT 13	
BX101540/c	
LOCUS	BX101540
DEFINITION	468 bp mRNA linear EST 06-FEB-2003
IMAGE:2456727,	NCI CGAP Gas4 Homo sapiens cDNA clone IMAGE998A166110 ; mRNA sequence.

ACCESSION	BX101540
VERSION	BX101540.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

TITLE
 JOURNAL
 COMMENT
 Reference
 Authors
 Eutheria: Metazoa: Chordata: Vertebrata: Euteleostomi:
 Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 1 (bases 1 to 468)
 Ebert, L., Hell, O., Hennig, S., Neubert, P., Parsch, E., Peters, M.,
 Radelfof, U., Schneider, D. and Korn, B.
 Human unigeneset - RZPD3
 Unpublished
 Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGP98A166110.
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No. 972)
<http://www.rzpd.de/CloneCards/cg1-b1/showlib.pl.cgi/response?libNo=972> contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heudenweg 6, D-14059 Berlin, Germany
Tel.: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD.
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13u, Primer sequence: CGTTGTAACGACGGCCACT.

FEATURES	Location/Qualifiers
source	1. .468

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998A166110 ; IMAGE:245672"
/tissue_type="poorly differentiated adenocarcinoma with
signal ring cell features"
/lab_host="DH10B"
/clone_idb="NCI CGAP Gas4"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

```

BASE COUNT	99 a	129 c	115 g	125 t
ORIGIN				
Query Match		23.1%;	Score 247.6;	DB 13;
Best Local Similarity		86.2%;	Pred.No. 1.3e+03;	Length 468;

	Matches	274;	Conservative	0;	Mismatches	44;	Indels	0;	Gaps	0;
QY	610	AATTTAAAGC	AATTCAGCCGCGGCCATGCTCA	CGCTGTAA	TCCAGACATT	TAGA	669			
Db	318	AAAGAAAAAGCT	GGCCCCAGCGGTGTGCTAC	CGCTTAAT	TCCAGACATT	TAGA	259			
QY	670	GGCGGAGGCGGGCGGATTCACCTGA	GGTGGAGTTGCGACACGAGCTTGGCA	CAATG	GTG	729				
Db	258	GGCGGAGGCGGGCGGATTC	CGGAGATCAGAGTTT	CGAGACGAGCTTGGCA	CAATG	GTG	199			
QY	730	AAACCTGTCTCTCA	AAAAAATATACAAAAAT	TAGCCGGGAGTGTG	ATGCTGTGAT	789				
Db	198	AAACCTGTCTCTCTCA	AAAAAATATACAAAAAT	TAGCTGTGCGTGTG	GTGCTGTGAT	139				
QY	790	CCGAGCTACTCGGAGGCTGAGG	CAGAGAAATCGCTTAA	ACC	CGGAGGCGGAGT	TGCA	849			
Db	138	CCGAGCTACTCTGGAAGGCTGAGG	CAGAGAAATCGCTTAA	ACC	CGGAGGCGGAGT	TGCA	79			
QY	850	GTGAGCCGAGATTCACAC	CGTCAGCTCAGCTGGGCGCA	AAAGACGAAAT	TTCGCTTAA	909				
Db	78	GTGAGCCGAGATTCACAC	CGTCAGCTCAGCTGGGCA	AAAGACGAGACTTCA	CTCA	19				
QY	910	AAAAATTAATTA	AAATAA	927						
Db	18	AAAGAAAAA	AAAAAAA	1						

RESULT 14	
AQ387027	
LOCUS	678 bp DNA linear
DEFINITION	GSS 21-MAY-1995
	Rpci11-153c12.TJ Rpci-11 Homo sapiens genomic clone Rpci11-153c12 genomic survey sequence.

ACCESSION	AQ387027
VERSION	AQ387027.1
KEYWORDS	GI:4358050
SOURCE	GSS.
	<i>Homo sapiens</i> (human)

ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE	1 (bases 1 to 6/8)
AUTHORS	Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building
JOURNAL	Unpublished
COMMENT	Other_GSSs: RPCI11-153C12..rv

Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genomics (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.

```

FEATURES
    source
        Location/Qualifiers
            1..678
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="GDB:7558427"
                /db_xref="taxon:9606"
                /clone="RPC1-11-153C12"
                /sex="Male"
                /cell_type="lymphocytes"
                /clone_lib="RPC1-11"
                /note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPC11 Human Male BAC Library"

```



```

: CURRENT APPLICATION NUMBER: US/09/641,638
: CURRENT FILING DATE: 2000-08-16
: PRIOR APPLICATION NUMBER: US 09/502,330
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: US 60/133,200
: PRIOR FILING DATE: 1999-05-07
: PRIOR APPLICATION NUMBER: US 09/275,267
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: US 60/119,917
: PRIOR FILING DATE: 1999-02-12
: NUMBER OF SEQ ID NOS: 1304
: SOFTWARE: Patent.pm
: SEQ ID NO 316
: LENGTH: 1001
: TYPE: DNA
: ORGANISM: Homo Sapiens
: FEATURE:
: NAME/KEY: allele
: LOCATION: 501
: OTHER INFORMATION: 12-361-320 : polymorphic base G or T
: NAME/KEY: misc_binding
: LOCATION: 481..500
: OTHER INFORMATION: 12-361-320.misl, potential
: NAME/KEY: misc_binding
: LOCATION: 502..520
: OTHER INFORMATION: 12-361-320.mis2, complement
: NAME/KEY: primer_bind
: LOCATION: 182..200
: OTHER INFORMATION: upstream amplification primer
: NAME/KEY: primer_bind
: LOCATION: 677..696
: OTHER INFORMATION: downstream amplification primer, complement
: NAME/KEY: misc_binding
: LOCATION: 489..513
: OTHER INFORMATION: 12-361-320 potential probe
: US-09-641-638-316

Query Match      44.7%: Score 22.8; DB 4; Length 1001;
Best Local Similarity 71.4%: Pred. No. 6.3;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      5 CCTTTAAGGCGTTGAATGTCGCACTGTCATGTGACT 46
DB      741 CCTTTTCGCTTCATTTCCCATCTGCAATGTGACATT 782

RESULT 3
: US-09-641-638-317
: Sequence 317, Application US/09641638
: Patent No. 6432648
: GENERAL INFORMATION:
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Bouguenelret, Lydie
: APPLICANT: Chumakov, Ilya
: APPLICANT: Cohen, Annick
: TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
: FILE REFERENCE: GENSET.05ICP1
: CURRENT APPLICATION NUMBER: US/09/641,638
: CURRENT FILING DATE: 2000-08-16
: PRIOR APPLICATION NUMBER: US 09/502,330
: PRIOR FILING DATE: 2000-02-11
: PRIOR APPLICATION NUMBER: US 60/133,200
: PRIOR FILING DATE: 1999-05-07
: PRIOR APPLICATION NUMBER: US 09/275,267
: PRIOR FILING DATE: 1999-03-23
: PRIOR APPLICATION NUMBER: US 60/119,917
: PRIOR FILING DATE: 1999-02-12
: NUMBER OF SEQ ID NOS: 1304
: SOFTWARE: Patent.pm
: SEQ ID NO 317
: LENGTH: 1001
: TYPE: DNA
```

```

: ORGANISM: Homo Sapiens
: FEATURE:
: NAME/KEY: allele
: LOCATION: 501
: OTHER INFORMATION: 12-361-388 : polymorphic base A or G
: NAME/KEY: misc_binding
: LOCATION: 482..500
: OTHER INFORMATION: 12-361-388.misl
: NAME/KEY: misc_binding
: LOCATION: 502..521
: OTHER INFORMATION: 12-361-388.mis2, potential complement
: NAME/KEY: primer_bind
: LOCATION: 114..132
: OTHER INFORMATION: upstream amplification primer
: NAME/KEY: primer_bind
: LOCATION: 609..628
: OTHER INFORMATION: downstream amplification primer, complement
: NAME/KEY: misc_binding
: LOCATION: 489..513
: OTHER INFORMATION: 12-361-388 potential probe
: US-09-641-638-317

Query Match      44.7%: Score 22.8; DB 4; Length 1001;
Best Local Similarity 71.4%: Pred. No. 6.3;
Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY      5 CCTTTAAGGCGTTGAATGTCGCACTGTCATGTGACT 46
DB      673 CCTTTTCGCTTCATTTCCCATCTGCAATGTGACATT 714

RESULT 4
: US-08-701-233B-2/C
: Sequence 2, Application US/08701233B
: Patent No. 5861308
: GENERAL INFORMATION:
: APPLICANT: Pfriundschuh, Michael; Renner, Christoph
: TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS ASSOCIATED WITH T CELL
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Felte & Lynch
: STREET: 805 Third Avenue
: CITY: New York City
: STATE: New York
: ZIP: 10022
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
: COMPUTER: IBM
: OPERATING SYSTEM: PC-DOS
: SOFTWARE: Wordperfect
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/701,233B
: FILING DATE: 21-August-1996
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Hanson, No. 5861308man D.
: REGISTRATION NUMBER: 30,946
: REFERENCE/DOCKET NUMBER: LUD 5440
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 688-9200
: TELEFAX: (212) 838-3884
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 380 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-701-233B-2

Query Match      42.0%: Score 21.4; DB 2; Length 380;
Best Local Similarity 63.3%: Pred. No. 17;
Matches 31; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
```


43 2 GCTGCTTTAAATGTTTGAATTTCTAAATGTTTATGTAACCTTTAA 50
 44 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 45 439 GCAATGCTTCTTGCTTTGATTTTAAATTTTAACTCTCTGAAATAA 252

RESULT 5

US-09-495-050A-156
 Sequence 156, Application US/09/45050A
 Patent No. 6492505

GENERAL INFORMATION:
 APPLICANT: Koopa, Peppy
 APPLICANT: Goetzel, Karl J.
 APPLICANT: Au-Yang, Janice
 TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
 FILE REFERENCE: EA-011408
 CURRENT APPLICATION NUMBER: US/09/45050A
 CURRENT FILING DATE: 2000-03-24
 PRIOR APPLICATION NUMBER: 60/2136,416
 PRIOR FILING DATE: February 1, 1999
 NUMBER OF SEQ ID NOS: 905
 SOFTWARE: Perl, Prostatm
 SEQ ID NO 156
 LENGTH: 2312
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURES:
 NAME/KEY: misc_feature
 OTHER INFORMATION: biotype ID No. 6492505 1687080C11
 US-09-495-050A-156

Query Match 42.0% Score 21.41 DB 4: Length 2412:
 Host Local Similarity 46.0% Prod. No. 28:
 Matches 31: Conservative 0: Mismatches 16: Indels 0: Gaps 0:

43 4 TCTTTTAAATGCTTGAATTTTAAATGTTATGTTATGTAACCTTAA 50
 44 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 45 1143 TCTTTTAAATGATGATTTTAAATTTAAATTTAAATTTGTTTAA 1189

RESULT 4

US-09-744-674-170
 Sequence 3, Application US/07/34674
 Patent No. 6498122

GENERAL INFORMATION:
 APPLICANT: WEL, Michael, et al.
 TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: ANTIBODIES THEREOF
 FILE REFERENCE: GLO-1004
 CURRENT APPLICATION NUMBER: US/07/34,674
 CURRENT FILING DATE: 2000-12-15
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 2020 nt
 TYPE: DNA
 ORGANISM: Human
 FEATURES:
 NAME/KEY: misc_feature
 TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
 TITLE OF INVENTION: ANTIBODIES THEREOF
 OTHER INFORMATION: A T17 of 3
 US-09-744-674-3

Query Match 42.0% Score 21.41 DB 4: Length 2020:
 Host Local Similarity 46.0% Prod. No. 92:
 Matches 31: Conservative 0: Mismatches 16: Indels 0: Gaps 0:

43 5 TCTTTTAAATGCTTGAATTTTAAATGTTATGTTATGTAACCTTAA 51
 44 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 45 106876 TATTTGAAATGATGATTTATTTTAAATGATGATTTAA 107830

RESULT 7
 US-09-186-276B-2670
 Sequence 26, Application US/09/186276B
 Patent No. 6484174

GENERAL INFORMATION:
 APPLICANT: Beeley, Philip
 APPLICANT: Delamater, Laura
 APPLICANT: Wyszok-Diller, Joanna
 APPLICANT: Malany, Jocelyn E.
 APPLICANT: psych, Leonard
 APPLICANT: Helanlita, Yrjo
 TITLE OF INVENTION: Scarecrow gene, Promoter and Uses Thereof
 FILE REFERENCE: 5914-075-999
 CURRENT APPLICATION NUMBER: US/09/186,276B
 CURRENT FILING DATE: 1998-11-05
 PRIOR APPLICATION NUMBER: 08/842,445
 PRIOR FILING DATE: 1997-04-24
 PRIOR APPLICATION NUMBER: 08/76,617
 PRIOR FILING DATE: 1996-04-26
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 26
 LENGTH: 1094
 TYPE: DNA
 ORGANISM: Zea mays
 US-09-186-276B-26

Query Match 41.6% Score 21.21 DB 4: Length 1094:
 Host Local Similarity 69.0% Prod. No. 27:
 Matches 29: Conservative 0: Mismatches 13: Indels 0: Gaps 0:

43 10 TAAAGCTTGAAGCTGCTGCACTGTCATGCTGATGATGATGATGAA 51
 44 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 45 927 TAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 966

RESULT 6

US-08-842-445-2570
 Sequence 25, Application US/08/842445A
 Patent No. 6441270

GENERAL INFORMATION:
 APPLICANT: Beeley et al.
 TITLE OF INVENTION: Scarecrow gene, Promoter and Uses
 TITLE OF INVENTION: 106904
 FILE REFERENCE: 5914-056-999
 CURRENT APPLICATION NUMBER: US/08/842,445A
 CURRENT FILING DATE: 1997-04-24
 EARLIER APPLICATION NUMBER: 08/76,617
 EARLIER FILING DATE: 1996-04-26
 NUMBER OF SEQ ID NOS: 79
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 26
 LENGTH: 1094
 TYPE: DNA
 ORGANISM: Plant
 US-08-842-445-25

Query Match 41.6% Score 21.21 DB 4: Length 1094:
 Host Local Similarity 69.0% Prod. No. 27:
 Matches 29: Conservative 0: Mismatches 13: Indels 0: Gaps 0:

43 10 TAAAGCTTGAAGCTGCTGCACTGTCATGCTGATGATGATGATGAA 51
 44 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT
 45 927 TAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 966

RESULT 3

US-09-186-276B-2670
 Sequence 26, Application US/09/186276B
 Patent No. 6484174

GENERAL INFORMATION:
 APPLICANT: Beeley et al.
 TITLE OF INVENTION: Scarecrow gene, Promoter and Uses Thereof

1 TITLE OF INVENTION: Thereof
2 FILE REFERENCE: 5914-074-999
3 CURRENT APPLICATION NUMBER: US/09/186,188B
4 CURRENT FILING DATE: 1998-11-05
5 PRIOR APPLICATION NUMBER: 08/842,445
6 PRIOR FILING DATE: 1997-04-24
7 PRIOR APPLICATION NUMBER: 08/638,617
8 PRIOR FILING DATE: 1996-04-26
9 NUMBER OF SEQ ID NOS: 79
10 SOFTWARE: FastSeq for Windows Version 3.0
11 SEQ ID NO 26
12 LENGTH: 1094
13 TYPE: DNA
14 ORGANISM: Plant
15 US-09-186-188B-26

Query Match 41.6%; Score 21.2; DB 4; Length 1094;
Best Local Similarity 69.0%; Pred. No. 27;
Matches 29; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 10 TTAAGGCTTGATGTCGCACTGTCATGTGACCTTAAG 51
DB 927 TTACGGTTTCCTTCTTCTACATCATTTGTACCTTACAG 886

RESULT 10
US-08-248-839C-9
Sequence 9, Application US/08248839C
Patent No. 5843702
GENERAL INFORMATION:
APPLICANT: McConnell, David
APPLICANT: Devine, Kevin
APPLICANT: O'Kane, Charles
TITLE OF INVENTION: A Gene Expression System
NUMBER OF SEQUENCES: 185
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 58437020 No. 5843702disk of No. 5843702th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,839C
FILING DATE: 25-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 3614,214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 180..518
US-08-248-839C-9

Query Match 41.6%; Score 21.2; DB 2; Length 1201;
Best Local Similarity 76.5%; Pred. No. 28;
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TGCTCCTTTAAGGCTTGATGTCGCACTGT 34
DB 64 TTCTCCTTTAACAATCATGTATGTTGAACACTGT 97

RESULT 11
US-09-080-855-17/c
Sequence 17, Application US/09080855A
Patent No. 6083721
GENERAL INFORMATION:
APPLICANT: Saras, Jan
APPLICANT: Franz, Petra
APPLICANT: Aspenstrm, Pontus
APPLICANT: Hellman, ulf
APPLICANT: Genez, Leonel Jorge
APPLICANT: Heldin, Carl-Henrik
TITLE OF INVENTION: PARC, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1
FILE REFERENCE: L0461/7030
CURRENT APPLICATION NUMBER: US/09/080,855A
CURRENT FILING DATE: 1998-05-18
EARLIER APPLICATION NUMBER: 08/805,583
EARLIER FILING DATE: 1997-02-25
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 458
TYPE: DNA
ORGANISM: Mus musculus
US-09-080-855-17

Query Match 41.2%; Score 21; DB 3; Length 458;
Best Local Similarity 73.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 TTTAAGGCTTGATGTCGCACTGTCATGTGTACA 44
DB 414 TTATGTGCTTGAGGCTGCAATTGTCTGTGTAGA 378

RESULT 12
US-09-566-076-17/c
Sequence 17, Application US/09566076
Patent No. 6475775
GENERAL INFORMATION:
APPLICANT: Saras, Jan
APPLICANT: Franz, Petra
APPLICANT: Aspenstrm, Pontus
APPLICANT: Hellman, ulf
APPLICANT: Genez, Leonel Jorge
APPLICANT: Heldin, Carl-Henrik
TITLE OF INVENTION: PARC, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1
FILE REFERENCE: L0461/7030
CURRENT APPLICATION NUMBER: US/09/566,076
CURRENT FILING DATE:
EARLIER APPLICATION NUMBER: 09/080,855
EARLIER FILING DATE: 1998-05-18
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 458
TYPE: DNA
ORGANISM: Mus musculus
US-09-566-076-17

Query Match 41.2%; Score 21; DB 4; Length 458;
Best Local Similarity 73.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 TTTAAGGCTTGATGTCGCACTGTCATGTGTACA 44
DB 414 TTATGTGCTTGAGGCTGCAATTGTCTGTGTAGA 378

RESULT 13

US-09-085-761A-45/-

Sequence 45, Application US/09085761A

Patent No. 6335178

GENERAL INFORMATION:

APPLICANT: Welner, Joel H.

APPLICANT: Turner, Raymond J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN

TITLE OF INVENTION: SECRETION

NUMBER OF SEQUENCES: 77

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/085,761A

FILING DATE: 28-MAY-1998

CLASSIFICATION: 405

ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Peter G.

REGISTRATION NUMBER: 32,847

REFERENCE/DOCKET NUMBER: DALB-03356

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8438

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 4120 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-09-085-761A-45

Query Match 41.2% Score 21; DB 4; Length 3120;

Best Local Similarity 56.7%; Prod. No. 4;

Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 5 GCCTTTAAAGGCTTGAATGCTGCAATGTCATGTCAGACTTAA 49

DB 1507 GGTATATAGGCTTGAATGTCATGTCAGACTTAACTTAA 1463

RESULT 14

US-09-053-197A-3/-

Sequence 3, Application US/09053197A

Patent No. 6022952

GENERAL INFORMATION:

APPLICANT: Welner, Joel H.

APPLICANT: Turner, Raymond J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN

TITLE OF INVENTION: SECRETION

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/053,197A

FILING DATE: 01-APR-1998

CLASSIFICATION: 405

ATTORNEY/AGENT INFORMATION:

NAME: MacKnight, Kamrin L.

REGISTRATION NUMBER: 38,230

REFERENCE/DOCKET NUMBER: DALB-03293

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8438

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 22108 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

US-09-053-197A-3

Query Match 41.2% Score 21; DB 3; Length 22108;

Best Local Similarity 56.7%; Prod. No. 7;

Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 5 GCCTTTAAAGGCTTGAATGCTGCAATGTCATGTCAGACTTAA 49

DB 17559 GGTATATAGGCTTGAATGTCATGTCAGACTTAACTTAA 17515

RESULT 15

US-09-085-761A-4/-

Sequence 3, Application US/09085761A

Patent No. 6335178

GENERAL INFORMATION:

APPLICANT: Welner, Joel H.

APPLICANT: Turner, Raymond J.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN

TITLE OF INVENTION: SECRETION

NUMBER OF SEQUENCES: 77

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/085,761A

FILING DATE: 28-MAY-1998

CLASSIFICATION: 405

ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Peter G.

REGISTRATION NUMBER: 32,847

REFERENCE/DOCKET NUMBER: DALB-03356

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8438

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 22108 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

US-09-085-761A-3

Query Match 41.2% Score 21; DB 4; Length 22108;

Best Local Similarity 66.7%; Pred. No. 73;
Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 5 CCTTTAAGGGCTTGATGTCTGCACTGTGATGTGACACTTAA 49
| | | | | | | | | | | | | | | | | | | | | |
Db 17559 CGTGATTAAGCGGTTGAGTATCTCTACAGACATGTTTACGGTTTA 17515

Search completed: August 20, 2003, 14:32:41
Job time : 8.26518 secs

GenCore version 5.1.6
Copyright (c) 1995 - 2003 CompuGen Ltd.

GM nucleic - nucleotide search, using SW model

Run on: August 20, 2003, 09:44:27 : Search time 20.0283 Seconds
(without alignments)
5873.827 Million cell updates/sec

Header: US-09-761-466-2

Perfect score: 51

Sequence: 1 ucgcgccttaagccttca.....tgacatgtatcactaaag 51

Scoring table: IDENTITY: 100

Gapop 100, Gapext 1.0

Searched: 2552755 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 9

Maximum DB seq length: 209000000

Post-processing: Minimum Match 9%

Listing first 45 summaries

Database:

N_Geneseq_1371013:*

1: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1990.DAT:*

11: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1991.DAT:*

12: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1992.DAT:*

13: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1993.DAT:*

14: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1994.DAT:*

15: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1995.DAT:*

16: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1996.DAT:*

17: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1997.DAT:*

18: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1998.DAT:*

19: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA1999.DAT:*

20: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2000.DAT:*

21: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2001.DAT:*

22: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2002.DAT:*

23: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2003.DAT:*

24: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2004.DAT:*

25: /SIDSL/ncdata/a/geneseq/geneseqn-emb1/NA2005.DAT:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length (b)	HL	Description
1	51	100.0	51	22	AAS09959
2	51	100.0	137.2	22	DNA encoding mcsx/
3	51	100.0	744.6	22	Genomic DNA [1] enc
4	23.2	46.3	445.24	23	Human prostate spe
5	23.2	45.5	4189	23	DNA encoding nucle
6	23.2	45.1	28	19	AAV2-7028
7	22.8	44.7	1761	21	AA57642
8	22.8	44.7	1001	21	AA57644

C	9	22.8	44.7	1194	18	AAV74433	Stat1[cytotoxic act
C	10	22.8	44.7	1341	24	AAH68333	Streptococcus Fc1Y
C	11	22.8	44.7	13206	24	ABL14494	Drosophila melanog
C	12	22.8	44.7	14209	24	ABL19943	Drosophila melanog
C	13	22.8	44.7	15500	22	AAH15943	Human nervous syst
C	14	22.8	44.7	15500	22	AAS66607	Human cardiovascular
C	15	22.8	44.7	15500	25	ABZ78784	Secreted protein 4
C	16	22.8	44.7	15500	25	AAH16648	Human secreted pro
C	17	22.8	44.7	15500	25	ABZ67378	Human secreted pro
C	18	22.8	44.7	165199	24	ABK34460	Human cDNA clone
C	19	22.6	44.3	257	25	ABX24402	Human 3T3 fibroste
C	20	22.6	44.3	257	25	ABX27148	Human 3T3 fibroste
C	21	22.6	44.3	790	20	AAH64979	Cort. plastic pro
C	22	22.4	43.9	653	24	ABK31119	Human prostate spe
C	23	22.4	43.9	1437	24	ABK31120	Human prostate spe
C	24	22.4	43.9	1620	22	AAK78421	Human immunoglob
C	25	22.4	43.9	2641	24	AAH21252	Arabis thapsus thall
C	26	22.2	43.5	742	24	ABN26597	Human cDNA clone
C	27	22.2	43.5	752	22	AAH95676	Human fibroblast
C	28	22.2	43.5	1265	20	AAZ33541	Human fibroblast
C	29	22.2	43.5	2583	24	ABK52446	cDNA clone of CAK1
C	30	22.2	43.5	3634	24	ABK83587	Human cDNA clone
C	31	22.2	43.5	3634	24	ABN75095	Gene #1593 used to
C	32	22.2	43.5	3635	24	ABV78105	Hypoxia-regulated
C	33	22.2	43.5	3635	24	ABK09401	Human cDNA encodin
C	34	22.2	43.5	4635	25	ACF50143	Human breast asso
C	35	22.2	43.5	5459	22	ABA20336	Human breast asso
C	36	22.2	43.5	8400	22	ABV16040	Human breast asso
C	37	22.2	43.1	512	21	AAH44473	Human fibroblast
C	38	22.2	43.1	1567	21	AAH37983	Arabis thapsus thall
C	39	22.2	43.1	2951	22	AAH17985	Human cDNA clone
C	40	22.2	43.1	4111	22	AAH76982	Human fibroblast
C	41	22.2	43.1	167343	24	ABH44405	Stomach cancer cel
C	42	22.2	43.1	167343	24	ABH44405	Stomach cancer cel
C	43	21.8	42.7	408	16	AAH12760	Human cDNA clone
C	44	21.8	42.7	1943	22	ABH46453	Human breast cell
C	45	21.8	42.7	1943	22	ABH46453	Human breast cell

ALIGNMENTS

RESULT 1

AA509959 standard; DNA: 51 bp.

AA509959:

24-OCT-2001 (first entry)

XX DNA encoding mcsx/NKX2.5 homology domain A2.

DE

XX

XX mcsx/NKX2.5; cardiac enhancer; cardiac cell; cardiomyocyte induction;

KW

KW Therapeutic; heart tissue; gene therapy; mouse; us.

XX

OS

OS Mus musculus.

XX

XX W20H151006-A2.

XX

XX 19-JUN-2001.

XX

XX 16 JAN 2001: 2001W0-US01511.

XX

XX 14 JAN 2000: 2000US-0176419.

XX

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX

XX Lee JW. Ezmo S.

XX

XX WPI: 2001-451809/48.

XX

XX Now cardiac specific cell enhancer elements, used to specifically

XX express gene in cardiac cell, as early marker of cardiomyocyte

PT induction, e.g. for optimizing cardiomyocyte induction -
XX
PS Claim 1; Fig 5A; 66pp; English.
XX
CC The sequence represents the coding sequence of cardiac enhancer
CC mcsx/Nkx2.5 homology domain A2. The nucleic acid is useful for
CC specifically expressing a gene in a cardiac cell, as an earlier marker
CC of cardiomyocyte induction, e.g. for optimizing cardiomyocyte induction.
CC Genes expressed in the cardiac cell-specific manner are useful for the
CC targeted expression of genes encoding therapeutic proteins for the
CC treatment of damaged heart tissue. Cardiac specific enhancer elements may
CC be used for gene therapy.
XX
SQ Sequence 51 BP; 12 A; 10 C; 11 G; 18 T; 0 other;
XX
Query Match 100.0%; Score 51; DB 22; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.1e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 51
DB 1 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 51
RESULT 2
AAS09960 AAS09960 standard; DNA; 1072 BP.
XX
AC AAS09960:
XX
DT 24-OCT-2001 (first entry)
XX
DE DNA encoding mcsx/Nkx2.5 homology domains A1 + A2, and intervening DNA.
XX
KM mcsx/Nkx2.5 homology domain; cardiac enhancer; cardiac cell;
KM cardiomyocyte induction; therapeutic; heart tissue; gene therapy;
KW mouse; ds.
XX
OS Mus musculus.
XX
PN WO200151006-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01511.
XX
PR 14-JAN-2000; 2000US-0176419.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Lee IW, Izumo S;
XX
DR WPI; 2001-451809/48.
XX
PT New cardiac specific cell enhancer elements, useful for specifically
PT expressing gene in cardiac cell, as earlier marker of cardiomyocyte
PT induction, e.g. for optimizing cardiomyocyte induction -
XX
PS Claim 1; Fig 5B; 66pp; English.
XX
CC The sequence represents the coding sequence of cardiac enhancer
CC mcsx/Nkx2.5 homology domains A1 + A2 and intervening sequence. The
CC nucleic acid is useful for specifically expressing a gene in a cardiac
CC cell, as an earlier marker of cardiomyocyte induction, e.g. for
CC optimising cardiomyocyte induction. Genes expressed in the cardiac cell-
CC specific manner are useful for the targeted expression of genes encoding
CC therapeutic proteins for the treatment of damaged heart tissue. Cardiac
CC specific enhancer elements may be used for gene therapy.
XX
SQ Sequence 1072 BP; 273 A; 290 C; 297 G; 212 T; 0 other;
XX
Query Match 100.0%; Score 51; DB 22; Length 1072;
Best Local Similarity 100.0%; Pred. No. 6.6e-10;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 51
DB 1022 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 1072
RESULT 3
AAS09961 AAS09961 standard; DNA; 7836 BP.
XX
AC AAS09961:
XX
DT 24-OCT-2001 (first entry)
XX
DE Genomic DNA #1 encoding human Csx/Nkx2.5.
XX
KM Csx/Nkx2.5; cardiac enhancer; cardiac cell; cardiomyocyte induction;
KM therapeutic; heart tissue; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO200151006-A2.
XX
PD 19-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US01511.
XX
PR 14-JAN-2000; 2000US-0176419.
XX
PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX
PI Lee IW, Izumo S;
XX
DR WPI; 2001-451809/48.
XX
PT New cardiac specific cell enhancer elements, useful for specifically
PT expressing gene in cardiac cell, as earlier marker of cardiomyocyte
PT induction, e.g. for optimizing cardiomyocyte induction -
XX
PS Disclosure; Fig 4A; 66pp; English.
XX
CC The sequence represents the genomic sequence #1 of human Csx/Nkx2.5. The
CC nucleic acid is useful for specifically expressing a gene in a cardiac
CC cell, as an earlier marker of cardiomyocyte induction, e.g. for
CC optimising cardiomyocyte induction. Genes expressed in the cardiac cell-
CC specific manner are useful for the targeted expression of genes encoding
CC therapeutic proteins for the treatment of damaged heart tissue. Cardiac
CC specific enhancer elements may be used for gene therapy.
XX
SQ Sequence 7836 BP; 2164 A; 1938 C; 1907 G; 1825 T; 2 other;
XX
Query Match 100.0%; Score 51; DB 22; Length 7836;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 51
DB 4609 TGCTCCTTTTAAGGGCTTGATGTCGCACTGCATGTCATGTAACCTTAAG 4659
RESULT 4
ABN87717 ABN87717 standard; cDNA; 865 BP.
XX
AC ABN87717:
XX
DT 08-AUG-2002 (first entry)
XX
DE Human prostate specific gene cDNA sequence SEQ ID NO:68.
XX
KM Human; prostate specific gene; prostate specific protein; PSG; PSP;
KW prostate cancer; gene; ss.

```

XX OS Homo sapiens.
XX
XX UN W0200236808-A2.
XX
XX PU 10-MAY-2002.
XX
XX PE 05-MAY-2001: 2001W-0847281.
XX
XX PE 04-MAY-2000: 2000US-2457430.
XX
XX (D1AD-) DIADEXUS INC.
XX
XX PU Sun Y, Reclon H, Cher S, Liu C.
XX
XX WP1: 2002-471506/50.
XX
XX DE New prostate-specific nucleic acids and polypeptides, useful for
PI identifying, diagnosing, monitoring, staging, imaging, and treating
PI prostate cancer and non-cancerous disease states in prostate tissue
XX
XX Claim 1: Page 187: 254pp: English.
XX
XX ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),
XX and ABN87492 to ABN87495 represent human prostate-specific proteins (II)
XX from the present invention. (I) and (II) have cytostatic activity. (I)
XX can be used in gene therapy. The prostate-specific nucleic acids,
XX polypeptides and compositions from the present invention can be used for
XX identifying, diagnosing, monitoring, staging, imaging, and treating
XX prostate cancer and non-cancerous disease states in prostate tissue; for
XX identifying prostate tissue; for monitoring, identifying and/or designing
XX agonists and antagonists of the polypeptides; in gene therapy; in
XX producing transgenic animals and cells; for producing engineered prostate
XX tissue for treatment and research; and as elements in an array or
XX computer program for pattern recognition of prostate disorders. The
XX nucleic acids may be used as hybridisation probes to detect, characterise
XX and quantify hybridised nucleic acids in, and isolate hybridising
XX nucleic acids from, both genomic and transcript-derived nucleic acid
XX samples.
XX
XX Sequence 865 BP: 150 A; 126 C; 130 G; 169 T; 250 other:
XX
XX Query Match 46.3%; Score 23.6; MB 24; Length 865;
XX Best local similarity 76.3%; Pred. No. 25;
XX Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
XX
XX QY 6 CTTTAAAGGCGCTGAATGCTGTGCACTGTCACTGTGAC 43
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 240 CTTTAAAGGAAATCAATTTCTTCACTTAAATGTAAC 277
XX
XX RESULT 5
XX AAS88398/c
XX ID AAS88398 standard; cDNA: 3480 BP.
XX
XX AC AAS88398:
XX
XX DT 14-FEB-2002 (first entry)
XX
XX DE DNA encoding novel human diagnostic protein #24202.
XX
XX KM Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX PU W0200175067-A2.
XX
XX PU 11-OCT-2001.
XX
XX PE 16-MAR-2001: 2001WO-0508641.
XX
XX PE 31-MAR-2000: 2000US-1540217.
XX

```

```

PR 2-AUG-2000: 2000US-0644167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PU Urmanac RT, Liu C, Fang Y1;
XX
XX DR WP1: 2001-639362/73.
XX
XX DR P-PSLB; ABG24211.
XX
XX
XX PI New isolated polynucleotide and encoded polypeptides, useful in
PI diagnostics, forensics, gene mapping, identification of mutations
PI responsible for genetic disorders or other traits and to assess
PI biodiversity
XX
XX Claim 1: SEQ ID No 24202: 103pp: English.
XX
XX DE The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences, (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from Wipo
XX at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 3480 BP: 890 A; 1003 C; 946 G; 640 T; 1 other:
XX
XX Query Match 45.5%; Score 23.2; DR 23; Length 3480;
XX Best local similarity 70.5%; Pred. No. 50;
XX Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
XX
XX QY 5 CTTTAAAGGCGCTGAATGCTGTGCACTGTGTAACACTTA 48
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 1821 CTTTGAAGGCGCTGAATGCTGTGCTCTTATATGAGGACTATA 1778
XX
XX RESULT 6
XX AAV26023
XX ID AAV26023 standard; DNA: 528 BP.
XX
XX AC AAV26023:
XX
XX DT 28-AUG-1998 (first entry)
XX
XX DE Bacteriophage DNA SEQ ID No:4-5.
XX
XX KM Bacteriophage; pathogen; Escherichia coli 0157; bacteriophage; food;
XX sterlisation; enteric haemorrhage; ds.
XX
XX OS Bacteriophage.
XX
XX PU W09808944-A1.
XX
XX PU 05-MAR-1998.
XX
XX PU 26-AUG-1997: 97WO-0P02957.
XX
XX PE 19-APR-1997: 97JP-0135716.
XX
XX PE 26-AUG-1996: 96JP-0261142.
XX
XX PE 14-APR-1997: 97JP-0140246.
XX

```

```

XX (BIOV-) BIO VENTURE BANK CO LTD.
XX
XX Nishikori K, Takahashi S;
XX
XX WPI; 1998-230262/20.
XX
XX Bacteriophage with high specificity for particular pathogens such as
XX Escherichia coli 0157 - is incorporated in bactericides for food
XX sterilisation
XX
XX Claim 3; Page 45; 54pp; Japanese.
XX
XX The present sequence represents a DNA sequence from a novel
XX bacteriophage which has high specificity for pathogenic bacteria,
XX especially for strains of Escherichia coli causing enteric haemorrhage,
XX such as E. coli 0157. The bacteriophage can be incorporated into
XX bio-bactericidal compositions. These can contain more than one
XX bacteriophage strain in order to have a bactericidal effect against more
XX than one pathogen at the same time. The bio-bactericide may also contain
XX an amino-acid or other material to control the pH to 6.5-7.5 and ensure
XX stability of the bacteriophage, such as glycine, arginine or lysine.
XX The bio-bactericidal composition can be used for treating food (such as
XX meat or fish) to prevent bacterial contamination, e.g. in fresh food or
XX in the kitchens of restaurants, schools and other institutions. It may
XX be formulated as a spray. It can also be used to sterilise working
XX surfaces, aprons and other clothing. The bio-bactericides are safe to
XX humans but very potent against pathogenic bacteria.
XX
XX Sequence 528 BP; 171 A; 89 C; 129 G; 137 T; 2 other;
XX
XX
XX Query Match 45.1%; Score 23; DB 19; Length 528;
XX Best Local Similarity 72.5%; Pred. No. 37;
XX Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX
XX 11 AAGGCTTGAAATGCTGCAACTGTCATGTACACTTAA 50
XX ||||| ||||| ||||| || ||||| ||
XX 312 AAGAGCTTGAAATGCTGCAAGGTACATGTCGACCTCANA 351
XX
XX
XX RESULT 7
XX AAC57682
XX ID AAC57682 standard; DNA: 1001 BP.
XX
XX AAC57682;
XX
XX 25-JAN-2001 (first entry)
XX
XX Arachidonic acid metabolism related genomic biallelic marker #316.
XX
XX Human: biallelic marker; arachidonic acid metabolism; genotyping;
XX detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;
XX single nucleotide polymorphism; hybridisation assay; sequencing assay;
XX specific amplification assay; identification; ERM; 12-LO-RBM;
XX eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.
XX
XX Homo sapiens.
XX
XX MO200047771-A2.
XX
XX 17-AUG-2000.
XX
XX 11-FEB-2000; 2000MO-IB00184.
XX
XX 12-FEB-1999; 99US-0119917.
XX 23-MAR-1999; 99US-0275267.
XX 07-MAY-1999; 99US-0133200.
XX
XX (GEST ) GENSET.
XX
XX Blumenfeld M, Bougueleret L, Chumakov I;
XX
XX WPI; 2000-571881/53.

```

```

XX
XX Novel biallelic markers useful for detecting conditions and genotypes
XX associated with arachidonic acid metabolism -
XX
XX Claim 13; Page 501-502; 802pp; English.
XX
XX The present invention describes polynucleotides including biallelic
XX markers derived from genes involved in arachidonic acid metabolism and
XX from genomic regions flanking those genes. Methods from the present
XX invention may be used to select individuals for clinical trials and
XX predict responses to treatment with drugs. The polynucleotides may be
XX used in hybridisation assays, sequencing assays and specific
XX amplification assays for identifying an eicosanoid-related biallelic
XX marker (ERM) or 12-LO-related biallelic marker, and for amplifying a
XX segment of nucleotides containing an ERM. The polynucleotides are
XX useful in diagnostic kits. The markers may be used to detect conditions
XX and genotypes associated with arachidonic acid metabolism. AAC5367 to
XX AAC58018 and AAB24019 and AAB24020 represent sequences used in the
XX exemplification of the present invention.
XX N.B. Polymorphic bases (single nucleotide polymorphisms also known as
XX SNPs) in the polynucleotide sequences from the present invention have
XX been given as their corresponding degenerate bases e.g. a polymorphic
XX base of C or T has been given as Y.
XX
XX Sequence 1001 BP; 187 A; 264 C; 272 G; 277 T; 1 other;
XX
XX
XX Query Match 44.7%; Score 22.8; DB 21; Length 1001;
XX Best Local Similarity 71.4%; Pred. No. 52;
XX Matches 30; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
XX
XX
XX 5 CCTTTTAAAGGCTTGAAATGCTGCAACTGTCATGTGACT 46
XX ||||| ||||| ||||| || ||||| |||||
XX 741 CCTTTTCTGCTTCATTCATTCCTCCATCTGCAATGTGACACT 782
XX
XX
XX RESULT 8
XX AAC57683
XX ID AAC57683 standard; DNA: 1001 BP.
XX
XX AAC57683;
XX
XX 25-JAN-2001 (first entry)
XX
XX Arachidonic acid metabolism related genomic biallelic marker #317.
XX
XX Human: biallelic marker; arachidonic acid metabolism; genotyping;
XX detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;
XX single nucleotide polymorphism; hybridisation assay; sequencing assay;
XX specific amplification assay; identification; ERM; 12-LO-RBM;
XX eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.
XX
XX Homo sapiens.
XX
XX WO200047771-A2.
XX
XX 17-AUG-2000.
XX
XX 11-FEB-2000; 2000MO-IB00184.
XX
XX 12-FEB-1999; 99US-0119917.
XX 23-MAR-1999; 99US-0275267.
XX 07-MAY-1999; 99US-0133200.
XX
XX (GEST ) GENSET.
XX
XX Blumenfeld M, Bougueleret L, Chumakov I;
XX
XX WPI; 2000-571881/53.
XX
XX Novel biallelic markers useful for detecting conditions and genotypes
XX associated with arachidonic acid metabolism -
XX
XX Claim 13; Page 502; 802pp; English.

```


XX the present invention describes polynucleotides including biallelic
 CC markers derived from genes involved in arachidonic acid metabolism and
 CC from genomic regions flanking those genes. Methods from the present
 CC invention may be used to select individuals for clinical trials and
 CC predict responses to treatment with drugs. The polynucleotides may be
 CC used in hybridization assays, sequencing assays and specific
 CC amplification assays for identifying an endonuclease-related biallelic
 CC marker (ERBM) or 12-10-related biallelic marker. The polynucleotides are
 CC segment of nucleotides containing an ERBM. The polynucleotides are
 CC useful in diagnostic kits. The markers may be used to detect conditions
 CC and genotypes associated with arachidonic acid metabolism. AAC57367 to
 CC AAC58018 and AAB24014 and AAB24020 represent sequences used in the
 CC exemplification of the present invention.
 CC N.B. Polymorphic bases (single nucleotide polymorphisms also known as
 CC SNPs) in the polynucleotide sequences from the present invention have
 CC been given as their corresponding degenerate bases e.g. a polymorphic
 CC base of C or T has been given as Y.

Sequence 1001 BP: 181 A: 279 C: 281 G: 278 T: 1 other;

Query Match 44.7%; Score 22.8; DB 21; Length 1001;
 Best Local Similarity 71.4%; Pred. No. 52;
 Matches 40; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

5 GCCTTAAAGGCTTGAAGTGTGCACTGCTGCTGACT 46
 11111 11111 11111 11111 11111 11111
 DB 673 GCCTTTCTGCTTGAATTTGGCATTTGTGTAATGTGACAT 714

RESULT 9

AAV74433/C
 ID AAV74433 standard; DNA: 1194 BP.

AAV74433;

16-MAR-1999 (first entry)

Staphylococcus aureus conidia SEQ ID #122.

Computer readable medium; vaccine; S. aureus infection; immunodetection;
 cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 skin infection; surgical wound infection; scalded skin syndrome;
 toxic shock syndrome; ds.

Staphylococcus aureus.

EP76519-A2.

40-JUL-1997.

07-JAN-1997; 97EP-010017.

05-JAN-1996; 960S-0009661.

(HUMA-) HUMAN GENOME SCI INC.

Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
 Rosen CA;

WPI: 1997-37942/45.

Polynucleotide(s) and proteins derived from Staphylococcus aureus
 stored on computer readable medium and used in the production of
 anti-S. aureus vaccines

Claim 1: Page 684; 327pp; English.

This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 of the invention. The DNA sequences are recorded on a computer readable
 medium, preferably selected from a floppy or hard disk, random access
 memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
 the S. aureus DNA sequences allows putative functions to be assigned so

CC that protein-encoding or regulatory regions of commercial, therapeutic or
 CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against S. aureus infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC S. aureus in a sample. S. aureus is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the S. aureus DNA sequences contained on the
 CC computer readable medium.

Sequence 1194 BP: 429 A: 158 C: 221 G: 383 T: 3 other;

Query Match 44.7%; Score 22.8; DB 16; Length 1194;
 Best Local Similarity 66.0%; Pred. No. 54;
 Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

2 GCCTTTTAAAGGCTTGAAGTGTGCACTGCTGCTGACT 51
 11111 11111 11111 11111 11111 11111
 DB 661 GCCTCTGCAATGCTGCACTGCAATGCAATGCAATGCAATGCAATGCAAT 612

RESULT 10

ABN68333/C
 ID ABN68333 standard; DNA: 1431 BP.

ABN68333;

01-JUL-2002 (first entry)

Streptococcus polynucleotide SEQ ID NO 4579.

Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 group A streptococcus; Streptococcus pyogenes; actinobacteria; group
 antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

Streptococcus pyogenes.

W020024771-A2.

02-MAY-2002.

29-06T-2001; 2001W-GB04789.

27-06T-2000; 2000GB-0026333.

24-NOV-2000; 2000GB-0026727.

07-MAR-2001; 2001GB-0005640.

(CHIR-) CHIRON SPA.

(GENO-) INST GENOMIC RES.

Telford J, Masigiani V, Margarit Ros YJ, Grandi G, Fraser G;
 Tettein H;

WPI: 2002-352536/38.

New Streptococcus protein for the treatment or prevention of infection
 or disease caused by Streptococcus bacteria, such as meningitis, and
 for detecting a compound that binds to the protein -

Claim 7: Page 3620; 4525pp; English.

The invention relates to a protein (ABP25413-ABP30895) from group B
 streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 the specification. The proteins have antibacterial and antiinflammatory
 activity. (1), nucleic acids encoding (1), ABN69044 ABN71526 and
 antibodies that bind (1) are used in the manufacture of pharmaceuticals for
 the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 1431 BP; 322 A; 298 C; 312 G; 499 T; 0 other;

Query Match 44.7%; Score 22.8; DB 24; Length 1431;
Best Local Similarity 66.0%; Pred. No. 57;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 2 GCTCTTTTAAAGGCGTTGAATGTCGCACTGTCATGTACACTTAAG 51
DB 573 GCCACCTGTAAAGCGCTTGAAGTCTTCAACGTCATGTTGAGCTGATAG 524

RESULT 11
ABL14494
ID ABL14494 standard; cDNA; 13206 BP.
XX
AC ABL14494;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37964.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PDB; ABB70391.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Claim 1; SEQ ID NO 37964; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 13206 BP; 4017 A; 2570 C; 2874 G; 3745 T; 0 other;

Query Match 44.7%; Score 22.8; DB 23; Length 13206;
Best Local Similarity 66.0%; Pred. No. 98;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1 TGCCTCTTTAAAGGCGTTGAATGTCGCACTGTCATGTACACTTAA 50
DB 9831 TGCCTCTTTCTTATAGATATGTCAGTAATATACATATGACACTTAA 9880

RESULT 12
ABL1940
ID ABL1940 standard; DNA; 14200 BP.
XX
AC ABL1940;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 11293.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
XX Claim 1; SEQ ID NO 11293; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 14200 BP; 4384 A; 2773 C; 3092 G; 3951 T; 0 other;

Query Match 44.7%; Score 22.8; DB 23; Length 14200;
Best Local Similarity 66.0%; Pred. No. 1e+02;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1 TGCCTCTTTAAAGGCGTTGAATGTCGCACTGTCATGTACACTTAA 50
DB 9200 TGCCTCTTTCTTATAGATATGTCAGTAATATACATATGACACTTAA 9249

RESULT 13
ABA15840
ID ABA15840 standard; DNA; 15500 BP.
XX
AC ABA15840;

PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX
PS Disclosure; SEQ ID NO 8171; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful.
CC In the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune deficiency, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC haemolytic anaemia, autoimmune arthritis, rheumatoid arthritis and ulcerative
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIFO at ftp.wifo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 15500 BP; 3660 A; 4223 C; 4207 G; 3410 T; 0 other;
Query Match 44.7%; Score 22.8; DB 22; Length 15500;
Best Local Similarity 66.0%; Pred. No. 1e+02;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
OY 1 TGCTCCTTTAAGCGCTGAATGTCGCACTGTCATGTGTACATTAA 50
Db 9272 TACTACTGTGAGATATTTAAAGATTTCCTACTGTCATATATTCATTAA 9321
RESULT 14
AAS36609/c
ID AAS36609 standard; DNA: 15500 BP.
XX
XX AAS36609;
AC
XX
DT 17-DEC-2001 (first entry)
XX
DE Human cardiovascular system antigen genomic DNA SEQ ID No 2109.
XX
XX
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotopic; dog;
KW antirheumatic; antiproliferative; cytostatic; cardiact; neuroprotective;
KW cerebroprotective; noctropic; antibacterial; vincinoid; fungicide; cancer;
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular infection; ds;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;
KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;

KW wound healing; skin aging; organ transplantation; tissue regeneration;
KW anti-infertility.
XX
XX
OS Homo sapiens.
XX
XX WO20015321-A2.
XX
XX
PD 02-AUG-2001.
XX
XX
PT 17-JAN-2001; 2001WO-US01340.
XX
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0199874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 18-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227099.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 12-SEP-2001; 2001US-0950082.
PR 12-SEP-2001; 2001US-0950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-040578/03.
XX
PT New human secreted proteins and nucleic acids, useful for detecting or
PT treating cancer or other hyperproliferative disorders, autoimmune
PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia -
XX
PS Disclosure; Page 1597-1601; 2474pp; English.
XX
CC AB273281-AB273697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC AB273698-AB274687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification, chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein genomic fragment referred to in the disclosure of the
CC invention.
XX
SQ Sequence 15500 BP; 3410 A; 4207 C; 4223 G; 3660 T; 0 other;

Query Match 44.7%; Score 22.8; DB 25; Length 15500;
Best Local Similarity 66.0%; Pred. No. 1e+02;
Matches 33; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 1 TGCTCCCTTTAAGCGCTTGAATGCTGCACTGTGATGTACACTTAA 50
| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6229 TACTACTGTGAAGATTTTAAGATTTCCTACTGTCATATATTCATTAA 6180

Search completed: August 20, 2003, 14:45:35
Job time : 22.0283 secs

Genotype version 5.1.6
Copyright (c) 1998 - 2003 CompuLink Ltd.

38 nucleotide probe sequence, using SW model

Run on: August 20, 2003, 09:44:27 : Search time 262.82 Seconds

(Without alignments)
7988.468 Million cell updates/sec

Title: us-09-761-466-2

Perfect score: 51

Sequence: 1 tgcctctttaaagactta.....tgcctgttctacttaag 51

Scoring table: IDENTITY_NWG

Gapop 10.0, Gapext 1.0

Searched: 2868711 seqs, 2045481366 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 6

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank

1: gb_ba*

2: gb_ba*

3: gb_ba*

4: gb_ba*

5: gb_ba*

6: gb_ba*

7: gb_ba*

8: gb_ba*

9: gb_ba*

10: gb_ba*

11: gb_ba*

12: gb_ba*

13: gb_ba*

14: gb_ba*

15: gb_ba*

16: gb_ba*

17: gb_ba*

18: gb_ba*

19: gb_ba*

20: gb_ba*

21: gb_ba*

22: gb_ba*

23: gb_ba*

24: gb_ba*

25: gb_ba*

26: gb_ba*

27: gb_ba*

28: gb_ba*

29: gb_ba*

30: gb_ba*

31: gb_ba*

32: gb_ba*

33: gb_ba*

34: gb_ba*

35: gb_ba*

36: gb_ba*

37: gb_ba*

38: gb_ba*

39: gb_ba*

40: gb_ba*

41: gb_ba*

score greater than or equal to the score of the result being filtered,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Result ID
1	51	100.0	81512	5	AC008412
2	51	100.0	137445	9	AF659424
3	39.8	78.0	198806	2	AF144621
4	30.2	59.2	171110	2	AF136699
5	30.2	59.2	241724	2	AF102990
6	29	56.9	193551	10	AF646055
7	27.8	54.5	98246	2	AF004394
8	27.8	54.5	142184	2	AF005775
9	27.8	54.5	159488	2	AF005823
10	27.2	53.3	66491	2	AF101861
11	27.2	53.3	190309	2	AF102750
12	27.2	52.9	184036	2	AF074174
13	26	51.0	101188	2	AF019524
14	26	51.0	183291	3	AF010603
15	26	51.0	255755	2	AF120618
16	26	51.0	264246	2	AF110408
17	26	51.0	271237	3	AF004521
18	25.8	50.6	228600	2	AF005093
19	25.8	50.6	248848	2	AF005103
20	25.4	49.8	240116	2	FX248432
21	25.2	49.4	130061	5	AF929217
22	25.2	49.4	136760	2	AF106175
23	25.2	49.4	221939	2	AF105299
24	25.2	49.4	256290	2	AF098754
25	25	49.0	68578	2	AF121329_3
26	25	49.0	80117	2	AF113350
27	25	49.0	96747	2	AF113350_3
28	25	49.0	106031	2	AF044168
29	25	49.0	110000	2	AF115160_2
30	25	49.0	124004	9	AF004947
31	25	49.0	134449	2	AF009049
32	25	49.0	140952	2	AF027182
33	25	49.0	143096	2	HS140012
34	25	49.0	147659	2	AF026312
35	25	49.0	162606	2	AF016348
36	25	49.0	166530	2	AF069259
37	25	49.0	166785	2	AF012568
38	25	49.0	173978	2	AF020649
39	25	49.0	186526	10	AF596083
40	25	49.0	191886	9	AF084200
41	25	49.0	197175	9	AF110292
42	25	49.0	204158	9	AF133863
43	25	49.0	216308	2	AF119542
44	25	49.0	239993	2	AF004312
45	24.8	48.6	171578	2	AF011731

ALIGNMENTS

RESULT 1	AC008412	81512 bp	DNA	Library	Ref. Seq. ID	Ref. Seq. Name
DEFINITION	Homo sapiens chromosome 5 clone, clone Z11114, complete sequence.					
ACCESSION	AC008412					
VERSION	AC008412.8					
KEYWORDS	HLA					
SOURCE	Homo sapiens (human)					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catartida; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 81512)					
AUTHORS	LVE Joint Genome Institute and Stanford Human Genome Center.					
TITLE	Direct Submission					
JOURNAL	Unpublished					

Prod. No. is the number of results predicted by chance to have a

```
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 81512)
DOE Joint Genome Institute.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 81512)
DOE Joint Genome Institute and Stanford Human Genome Center.
DOE Joint Genome Institute
Submitted (01-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
5 (bases 1 to 81512)
DOE Joint Genome Institute and Stanford Human Genome Center.
DOE Joint Genome Institute
Submitted (28-JUN-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 28, 2003 this sequence version replaced gi:20376952.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40.99.3% of Sequence:
Estimated Total Number of Errors is 0.4.
NOTE: This insert is not the entire sequence of the clone (entire
sequence is 139,581). It is clipped at the overlaps with AC008378
and AC106731. The number of bases overlapped with AC008378 is 4442
bps and with AC106731 is 15413 bps.
Location/Qualifiers
1. 81512
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-281H14"
BASE COUNT
20011 a 19534 c 19980 g 21987 t
ORIGIN
Query Match 100.0%; Score 51; DB 9; Length 81512;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGCTCCTTTAAGGGCTGAATGTCGCACTGTCATGTGACCTTAAG 51
Db 53141 TGCTCCTTTAAGGGCTGAATGTCGCACTGTCATGTGACCTTAAG 53091
RESULT 2
AL669924/c 137345 bp DNA linear PRI 24-APR-2002
LOCUS
DEFINITION
Human DNA sequence from clone Xkbac-2L19 on chromosome 6, complete
sequence.
ACCESSION
AL669924
VERSION
AL669924.9 GI:20068712
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
Organism
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19572797.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
COMMENT
```

```
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/WormPEP
from a CHORI-501 human bac - PGF cell line library VECTOR:
PTARBA2.1
This sequence was generated from part of bacterial clone contigs
constructed by the MHC Haplotype Consortium and collaborators.
Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC.
Location/Qualifiers
1. 137345
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="Xkbac-2L19"
/clone_lib="CHORI-501"
BASE COUNT
33377 a 33207 c 33325 g 37436 t
ORIGIN
Query Match 100.0%; Score 51; DB 9; Length 137345;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TGCTCCTTTAAGGGCTGAATGTCGCACTGTCATGTGACCTTAAG 51
Db 53367 TGCTCCTTTAAGGGCTGAATGTCGCACTGTCATGTGACCTTAAG 53317
RESULT 3
AC144621 158806 bp DNA linear HTG 06-MAY-2003
LOCUS
DEFINITION
Mus musculus chromosome UNK clone RP24-335114, WORKING DRAFT
SEQUENCE, 69 unordered pieces.
ACCESSION
AC144621
VERSION
AC144621.1 GI:30387755
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
Mus musculus (house mouse)
Organism
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (06-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (06-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
Project Information
-----
Center project name: M.BB0335114
-----
Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
```


Chemistry: Dye-Primer ET: 0% of reads
Chemistry: Dye-Terminator Big Dye: 100% of reads
Assembly program: Phrap version 0.990319
Consensus quality: 143426 bases at least Q40
Consensus quality: 146442 bases at least Q30
Consensus quality: 149961 bases at least Q20
Insert size: 417000; agarose-ip
Insert size: 153245; sum-of-contigs
Quality coverage: 1.91 in Q20 bases; sum-of-contigs
Quality coverage: 2.65 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 69 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1262: contig of 1262 bp in length
* 1263 1362: gap of unknown length
* 1363 2803: contig of 1447 bp in length
* 2810 2909: gap of unknown length
* 2910 4113: contig of 1204 bp in length
* 4114 4213: gap of unknown length
* 4214 5492: contig of 1279 bp in length
* 5493 5593: gap of unknown length
* 5593 5902: contig of 1310 bp in length
* 5903 7002: gap of unknown length
* 7003 8255: contig of 1253 bp in length
* 8256 8356: gap of unknown length
* 8357 9423: contig of 1068 bp in length
* 9424 9524: gap of unknown length
* 9524 10777: contig of 1254 bp in length
* 10778 10878: gap of unknown length
* 10879 12023: contig of 1146 bp in length
* 12024 12123: gap of unknown length
* 12124 13409: contig of 1286 bp in length
* 13410 13509: gap of unknown length
* 13510 15253: contig of 1744 bp in length
* 15254 15353: gap of unknown length
* 15354 16449: contig of 1296 bp in length
* 16450 16750: gap of unknown length
* 16751 18229: contig of 1480 bp in length
* 18230 18330: gap of unknown length
* 18331 19562: contig of 1331 bp in length
* 19563 19762: gap of unknown length
* 19763 21025: contig of 1262 bp in length
* 21026 21125: gap of unknown length
* 21126 22467: contig of 1343 bp in length
* 22468 22567: gap of unknown length
* 22568 23415: contig of 1048 bp in length
* 23416 23716: gap of unknown length
* 23717 25651: contig of 1936 bp in length
* 25652 25752: gap of unknown length
* 25753 27281: contig of 1530 bp in length
* 27282 27381: gap of unknown length
* 27382 29420: contig of 2039 bp in length
* 29421 29520: gap of unknown length
* 29521 40901: contig of 1381 bp in length
* 40902 41301: gap of unknown length
* 41302 42159: contig of 1158 bp in length
* 42160 42260: gap of unknown length
* 42261 43886: contig of 1528 bp in length
* 43887 43986: gap of unknown length
* 43987 45134: contig of 1247 bp in length
* 45135 45235: gap of unknown length
* 45236 46486: contig of 1254 bp in length
* 46487 46587: gap of unknown length
* 46588 48485: contig of 1799 bp in length
* 48486 48586: gap of unknown length
* 48587 49714: contig of 1229 bp in length
* 49715 49814: gap of unknown length

39815 41078: contig of 1254 bp in length
41079 41178: gap of unknown length
41179 42274: contig of 1095 bp in length
42275 42374: gap of unknown length
42375 44223: gap of unknown length
44224 45908: contig of 1665 bp in length
45909 46009: gap of unknown length
46010 47545: contig of 1537 bp in length
47546 47645: gap of unknown length
47646 50012: contig of 2367 bp in length
50013 50112: gap of unknown length
50113 51978: contig of 1866 bp in length
51979 52078: gap of unknown length
52079 54113: contig of 2035 bp in length
54114 55586: contig of 1374 bp in length
55587 55687: gap of unknown length
55688 57405: contig of 1719 bp in length
57406 57505: gap of unknown length
57506 58790: contig of 1281 bp in length
58791 58891: gap of unknown length
58892 60454: gap of unknown length
60455 62139: contig of 1596 bp in length
62140 62250: gap of unknown length
62251 63911: contig of 1661 bp in length
63912 65910: gap of unknown length
65911 66010: contig of 1900 bp in length
66011 67754: contig of 1744 bp in length
67755 67855: gap of unknown length
67856 69859: contig of 2005 bp in length
69860 71962: gap of unknown length
71963 72063: contig of 2003 bp in length
72064 74371: gap of unknown length
74372 74471: gap of unknown length
74472 76439: contig of 2028 bp in length
76440 76500: gap of unknown length
76501 78329: contig of 1730 bp in length
78330 78429: gap of unknown length
78430 81033: contig of 2604 bp in length
81034 81133: gap of unknown length
81134 82451: contig of 1317 bp in length
82452 82551: gap of unknown length
82552 85881: contig of 3330 bp in length
85882 85981: gap of unknown length
85982 89960: contig of 3979 bp in length
89961 90060: gap of unknown length
90061 92409: contig of 2440 bp in length
92410 92509: gap of unknown length
92510 95373: gap of unknown length
95374 95472: gap of unknown length
95473 98796: contig of 3324 bp in length
98797 98896: gap of unknown length
98897 101619: contig of 2724 bp in length
101620 101720: gap of unknown length
101721 105463: contig of 3744 bp in length
105464 105564: gap of unknown length
105565 108521: contig of 2957 bp in length
108522 108620: gap of unknown length
108621 111714: contig of 3094 bp in length
111715 111815: gap of unknown length
111816 115701: contig of 3887 bp in length
115702 115801: gap of unknown length
115802 117201: contig of 1919 bp in length
117202 117299: gap of unknown length
117300 121279: contig of 4459 bp in length
121280 121379: gap of unknown length
121380 125489: contig of 4110 bp in length
125490 125589: gap of unknown length
125590 128497: contig of 3308 bp in length

```

* 128898 12897: gap of unknown length
* 128998 13134: contig of 4137 bp in length
* 133135 13324: gap of unknown length
* 133235 137402: contig of 4168 bp in length
* 137403 137502: gap of unknown length
* 137503 142642: contig of 5140 bp in length
* 142643 142742: gap of unknown length
* 142743 148552: contig of 5810 bp in length
* 148553 148653: gap of unknown length
* 148653 158806: contig of 10154 bp in length.
Location/Qualifiers
source
  1. 158806
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /chromosome="X"
  /clone="RP24-335114"
  1. 1262
  /note="assembly_name:Contig10"
misc_feature
  1363. 2809
  /note="assembly_name:Contig22"
misc_feature
  2910. 4113
  /note="assembly_name:Contig34"
Query Match
  78.0%; Score 39.8; DB 2; Length 158806;
Best Local Similarity 86.3%; Pred. No. 5.7e-05;
Matches 44; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 TGCCTCTTTTAAAGGCTGATGCTGCAACGTCATGTCATGTCACCTTAAG 51
    |||||
Db 139093 TGCCTCTTTTAAAGGCTGATGCTGCAACGCTCATGTCACCTTAAG 139143

RESULT 4
LOCUS AC119699/c 171110 bp DNA linear HTG 13-NOV-2002
DEFINITION Rattus norvegicus clone CH230-466D21, WORKING DRAFT SEQUENCE, 5
unordered pieces.
ACCESSION AC119699.8 GI:24941740
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 171110)
AUTHORS Muzny,D.,Marle,,Metzker,M.,Lee,,Abramson,S.,Adams,C.,Alder,J.,
Allen,C.,Allen,H.,Alsbrooks,S.,Amin,A.,Angulano,D.,
Anyadebech,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
Baldwin,D.,Bandaraike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
Biswal,D.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderon,E.,
Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,
Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Diya,K.,
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Eaves,K.,
Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
Fraser,C.M.,Gabisi,A.,Ganta,R.,Grady,M.,Guerra,M.,Guevara,W.,
Gebregeorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,M.,Guevara,W.,
Gunnarsson,P.,Haaland,W.,Hamill,C.,Hamilton,C.,Hamilton,K.,
Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
Hollins,B.,Howells,S.,Huijck,S.,Hume,J.,Idlebird,D.,Jackson,A.,
Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,
Karpethy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kovar,C.,
Kowis,C.,Kraft,C.L.,Lebow,H.,Levan,D.,Lewis,L.,Li,Z.,Liu,J.,
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
Lorensuhewa,L.,Louiseged,H.,Lozado,R.J.,Lu,X.,Ma,J.,
Maheshwari,M.,Mahindartne,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mangum,B., Mapue,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nall,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwokilech,O., Okwouon,G., Olarnunagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Pioppe,F., Polindexter,A., Popovic,D., Plims,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Reiter,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Silter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,D.,
Steinle,M., Strong,R., Sutton,A., Sytek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczkyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,Y.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Weidenhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 171110)
Morley,K.C.
Direct Submission
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 171110)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 13, 2002 this sequence version replaced gi:23908283.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold and separated
individual sequence scaffolds are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVQK
Center clone name: CH230-466D21
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 159707 bases at least Q40
Consensus quality: 161692 bases at least Q30
Consensus quality: 163017 bases at least Q20
Estimated insert size: 161507; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 5 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.


```

* (see http://www.bjsgc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  1 241724: contig of 241724 bp in length.
FEATURES
  source
    1..241724
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-53024"
  misc_feature
    1..2211
    /note="wgs_contig"
BASE COUNT    60533 a 53609 c 56345 g 67996 t 3241 others
ORIGIN

```

```

Query Match      59.2%; Score 30.2; DB 2; Length 241724;
Best Local Similarity 74.3%; Pred. No. 0.38;
Matches 38; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

OY      1 TGCTCCTTTAAGGCTTGATGTCGACACTGCATGTGACACTTAAG 51
        |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      174895 TGCTCCTTTCAAGGCTTGAGTGGCTGCACCTTCTTGTCACCTTAAG 174845

```

```

RESULT 6
LOCUS    AL646055      193551 bp    DNA    linear    ROD 14-MAR-2002
DEFINITION
  Mouse DNA sequence from clone RP23-382C18 on chromosome 11,
  complete sequence.
ACCESSION
  AL646055
VERSION   AL646055.10  GI:19031711
KEYWORDS   HTG.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 193551)
REFERENCE
  AUTHORS   Kay, M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-MAR-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
            On Feb 28, 2002 this sequence version replaced gi:18476931.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em, EMBL; Sw,
            SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-382C18 is
            from the RPO1-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACE3.6
            This sequence is the entire insert of clone RP23-382C18 The true

```

```

FEATURES
  source
    1..193551
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="11"
    /clone="RP23-382C18"
    /clone_1fb="RPO1-23"
    /clone_1lb="RPO1-23"
    complement(186114..86118)
    /note="Sequence from uni-directional primer reads and dGTP
    big dye terminator reads only."
  misc_feature
    160975
    /note="Tandem repeat. Forced join. Assembly confirmed by
    restriction digest data."
BASE COUNT    57068 a 40716 c 39095 g 56672 t
ORIGIN

```

```

Query Match      56.9%; Score 29; DB 10; Length 193551;
Best Local Similarity 77.8%; Pred. No. 1.1;
Matches 35; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

OY      4 TCCTTTAAGGCTTGATGTCGACACTGCATGTGACACTTA 48
        |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      60646 TCCAGGATAGGCTTGATGTCGACACTGTATGTCCTTA 60602

```

```

RESULT 7
LOCUS    AP004394      98286 bp    DNA    linear    HTG 21-MAR-2002
DEFINITION
  Oryza sativa (japonica cultivar-group) chromosome 2 clone
  OJ1118_C06, *** SEQUENCING IN PROGRESS ***.
ACCESSION
  AP004394
VERSION   AP004394.1  GI:17402600
KEYWORDS   HTG; PHASE2.
SOURCE     Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
  1
  AUTHORS   Sasaki, T., Matsumoto, T. and Yamamoto, K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:OJ1118_C06
            Published Only in Database (2001)
            2 (bases 1 to 98286)
  JOURNAL   Submitted (05-DEC-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program, Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
  JOURNAL   The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
            NOTE: It currently consists of 1 contigs. Gaps between the contigs
            are represented as runs of N. The order of the pieces is believed
            to be correct as given, however the sizes of the gaps between them
            are based on estimates that have provided by the submitter. This
            sequence will be replaced by the finished sequence as soon as it is
            available and the accession number will be preserved.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.

```

```

FEATURES
  source
    1..98286
    /organism="Oryza sativa (japonica cultivar-group)"
    /mol_type="genomic DNA"
    /cultivar="Nipponbare"
    /db_xref="taxon:39947"
    /chromosome="2"
    /clone="OJ1118_C06"

```

BASE COUNT 26374 a 2329 c 22929 g 25741 t 213 others
 ORIGIN
 Query Match 54.5%: Score 27.8; DB 2; Length 96286;
 Best Local Similarity 74.5%: Prod. No. 3.5;
 Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 3 GCGCTTTAAGAGCTTGAAGTCTTACAGTCGATGCTACACTTAA 49
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 95691 GCGATTTCACACATATAATATCTTATCTGTCGATGACATTTAA 85737
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 8
 AP005823
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2, clone
 OS:INB0005A04. *** SEQUENCE IN PROGRESS ***
 ACCESSION
 VERSION AP005775.1 GI:21491627
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; Oryza.
 REFERENCE
 AUTHORS Sasaki, T., Matsumoto, I. and Katayose, Y.
 TITLE Oryza sativa nippondare(DAS) genomic DNA, chromosome 2, BAC
 clone:OS:INB0005A04
 JOURNAL Published only in Database (2002)
 REFERENCE
 AUTHORS Sasaki, T., Matsumoto, I. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 FEATURES
 SOURCE
 1..142184
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /submitter="Nippondare"
 /db_xref="taxon:3947"
 /chromosome="2"
 /clone="OS:INB0005A04"
 BASE COUNT 39311 a 31251 c 32076 g 39272 t 254 others
 ORIGIN
 Query Match 54.5%: Score 27.8; DB 2; Length 142184;
 Best Local Similarity 74.5%: Prod. No. 3.5;
 Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 3 GCGCTTTAAGAGCTTGAAGTCTTACAGTCGATGCTACACTTAA 49
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 37488 GCGATTTCACACATATAATATCTTATCTGTCGATGACATTTAA 37534
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 4
 AP005823
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2, clone p563F07,
 *** SEQUENCE IN PROGRESS ***

ACCESSION AP005823
 VERSION AP005823.1 GI:23821521
 KEYWORDS
 HTG: HTGS_PHASE2.
 Oryza sativa (japonica cultivar-group)
 SOURCE
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzoae; Oryza.
 REFERENCE
 1
 AUTHORS Sasaki, T., Matsumoto, I. and Katayose, Y.
 TITLE Oryza sativa nippondare(DAS) genomic DNA, chromosome 2, BAC
 clone:p063F07
 JOURNAL Published only in Database (2002)
 REFERENCE
 2 (bases 1 to 159985)
 AUTHORS Sasaki, T., Matsumoto, I. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-OCT-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 FEATURES
 SOURCE
 1..159985
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /submitter="Nippondare"
 /db_xref="taxon:3947"
 /chromosome="2"
 /clone="p063F07"
 BASE COUNT 42069 a 36621 c 37790 g 43207 t 301 others
 ORIGIN
 Query Match 54.5%: Score 27.8; DB 2; Length 159985;
 Best Local Similarity 74.5%: Prod. No. 3.5;
 Matches 35; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 3 GCGCTTTAAGAGCTTGAAGTCTTACAGTCGATGCTACACTTAA 49
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 4671 GCGATTTCACACATATAATATCTTATCTGTCGATGACATTTAA 4676
 ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 10
 AP010461
 LOCUS
 DEFINITION Mus musculus clone KP23 114K19, LOW PASS SEQUENCE SAMPLING.
 ACCESSION
 VERSION AC010361
 KEYWORDS
 HTG: HTGS_PHASE0.
 SOURCE
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 66891)
 AUTHORS Birren, B., Linton, J., Nussbaum, C., and Lander, E.
 TITLE Mus musculus, clone KP23-114K19
 JOURNAL Unpublished
 2 (bases 1 to 66891)
 AUTHORS Birren, B., Linton, J., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barina, N., Bastien, V., Boduslavsky, I., Bookshaiter, R.,
 Brown, A., Camarata, J., Campotien, A., Chang, J., Chazotte, R.,
 Choe, P., Choe, Y., Colaninno, M., Col, J., S., Collymore, A., Cox, K. A.,
 Cooke, P., DeAngelis, K., Dwyer, K., Elia, J., S., Farrow, S.,
 Ferreira, P., Fitzhugh, W., Gao, J., Galante, J., Gardiner, S.,

TITLE
JOURNAL
COMMENT

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Huime,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,
Lamares,R., Landers,T., Lehoccky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Melrim,J.,
Meneu,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C.,
Nordh,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollard,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Slit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L16535
Center Clone name: 113_K_19

* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 653: contig of 653 bp in length
* 654 753: gap of 100 bp
* 754 1469: contig of 716 bp in length
* 1470 1569: gap of 100 bp
* 1570 2270: contig of 701 bp in length
* 2271 2370: gap of 100 bp
* 2371 3061: contig of 691 bp in length
* 3062 3161: gap of 100 bp
* 3162 3877: contig of 716 bp in length
* 3878 3977: gap of 100 bp
* 3978 4686: contig of 709 bp in length
* 4687 4786: gap of 100 bp
* 4787 5500: contig of 714 bp in length
* 5501 5600: gap of 100 bp
* 5601 6315: contig of 715 bp in length
* 6316 6415: gap of 100 bp
* 6416 7136: contig of 721 bp in length
* 7137 7236: gap of 100 bp
* 7237 7944: contig of 708 bp in length
* 7945 8044: gap of 100 bp
* 8045 8752: contig of 708 bp in length
* 8753 8852: gap of 100 bp
* 8853 9564: contig of 712 bp in length
* 9565 9664: gap of 100 bp
* 9665 10375: contig of 711 bp in length
* 10376 10475: gap of 100 bp
* 10476 11184: contig of 709 bp in length
* 11185 11284: gap of 100 bp
* 11285 11999: contig of 715 bp in length
* 12000 12099: gap of 100 bp
* 12100 12774: contig of 675 bp in length

12775 12874: gap of 100 bp
12875 13593: contig of 719 bp in length
13594 13694: gap of 100 bp
13694 14400: contig of 707 bp in length
14401 14500: gap of 100 bp
14501 15216: contig of 716 bp in length
15217 15317: gap of 100 bp
15317 16009: contig of 693 bp in length
16010 16109: gap of 100 bp
16110 16806: contig of 697 bp in length
16807 16806: gap of 100 bp
16807 17624: contig of 718 bp in length
17625 17724: gap of 100 bp
17725 18432: contig of 708 bp in length
18433 18532: gap of 100 bp
18533 19230: contig of 698 bp in length
19231 19330: gap of 100 bp
19331 20053: contig of 723 bp in length
20054 20153: gap of 100 bp
20154 20853: contig of 700 bp in length
20854 20953: gap of 100 bp
20954 21665: contig of 712 bp in length
21666 21765: gap of 100 bp
21766 22480: contig of 715 bp in length
22481 22580: gap of 100 bp
22581 23294: contig of 714 bp in length
23295 23394: gap of 100 bp
23395 24094: contig of 700 bp in length
24095 24194: gap of 100 bp
24195 24897: contig of 703 bp in length
24898 24997: gap of 100 bp
24998 25713: contig of 716 bp in length
25714 25813: gap of 100 bp
25814 26529: contig of 716 bp in length
26530 26629: gap of 100 bp
26630 27345: contig of 716 bp in length
27346 27445: gap of 100 bp
27446 28166: contig of 721 bp in length
28167 28266: gap of 100 bp
28267 28984: contig of 718 bp in length
28985 29084: gap of 100 bp
29085 29769: contig of 685 bp in length
29770 29869: gap of 100 bp
29870 30507: contig of 638 bp in length
30508 30607: gap of 100 bp
30608 31327: contig of 720 bp in length
31328 31427: gap of 100 bp
31428 32151: contig of 724 bp in length
32152 32251: gap of 100 bp
32252 32945: contig of 694 bp in length
32946 33045: gap of 100 bp
33046 33760: contig of 715 bp in length
33761 33860: gap of 100 bp
33861 34571: contig of 711 bp in length
34572 34671: gap of 100 bp
34672 35387: contig of 716 bp in length
35388 35487: gap of 100 bp
35488 36185: contig of 698 bp in length
36186 36285: gap of 100 bp
36286 37003: contig of 718 bp in length
37004 37103: gap of 100 bp
37104 37804: contig of 701 bp in length
37805 37904: gap of 100 bp
37905 38619: contig of 715 bp in length
38620 38719: gap of 100 bp
38720 39423: contig of 704 bp in length
39424 39523: gap of 100 bp
39524 40228: contig of 705 bp in length
40229 41042: gap of 100 bp
41043 41142: contig of 714 bp in length
41143 41844: gap of 100 bp
41845 41944: gap of 100 bp

*	4194	contig of 717 bp	in length
*	42662	42761: gap of 100 bp	
*	42762	43475: contig of 714 bp	in length
*	43476	43575: gap of 100 bp	
*	43576	44285: contig of 710 bp	in length
*	44286	44385: gap of 100 bp	
*	44386	45098: contig of 713 bp	in length
*	45099	45198: gap of 100 bp	
*	45199	45893: contig of 695 bp	in length
*	45894	45993: gap of 100 bp	
*	45994	46717: contig of 724 bp	in length
*	46718	46817: gap of 100 bp	
*	46818	47492: contig of 675 bp	in length
*	47493	47592: gap of 100 bp	
*	47593	48407: contig of 715 bp	in length
*	48408	48407: gap of 100 bp	
*	48408	49123: contig of 716 bp	in length
*	49124	49223: gap of 100 bp	
*	49224	49320: contig of 697 bp	in length
*	49321	50023: gap of 100 bp	
*	50021	50738: contig of 718 bp	in length
*	50739	50838: gap of 100 bp	
*	50839	51547: contig of 709 bp	in length
*	51548	51647: gap of 100 bp	
*	51648	52344: contig of 697 bp	in length
*	52345	52444: gap of 100 bp	
*	52445	53162: contig of 718 bp	in length
*	53163	53262: gap of 100 bp	
*	53263	53974: contig of 712 bp	in length
*	53975	54074: gap of 100 bp	
*	54075	54780: contig of 705 bp	in length
*	54781	54880: gap of 100 bp	
*	54881	55579: contig of 699 bp	in length
*	55580	55679: gap of 100 bp	

Query Match	53.38;	Score 27.2;	DB 2:	Length 66891;
Best Local Similarity	72.98;	Pctid No. 6.2;		
Matches 35;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;
QY	3	GTCTTTTAAAGGTTGGATTCGTGCACATCATGTACTATAA	50	
Tb	65669	GTCTTTTCAACAATTGAAGAGTGCGATCTGCGACAGCATTCTTAA	65622	

RESULT 1:	
AC102750/c	19,809 bp linear HTG-21-0C1-202
LENGTHS	
DEFINITION	Mus musculus clone RP24-1A1H15, WORKING DRAFT SEQUENCE, 4 unordered pieces.
ACCESSION	AC102750
VERSION	AC102750.3 gi:24182059
KEYWORDS	HIG: HIGS-PHASE1; HIGS-DRAFT; HIGS-FILLTOP;
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Epkarytzka, M.; Zozom; Chordata: Cephalata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus. 1 (bases 1 to 19,809)
AUTHORS	Birren, B., Nisbaum, C., and Lander, E.
TITLE	Mus musculus, clone RP24-1A1H15
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 19,809)
AUTHORS	Birren, B., Illion, L., Nisbaum, C., Lander, E., Ali, A., Allen, N.,

Anderson, S., Samardja, N., Bastian, V., Boguslavsky, I., Bonkhallier, B., Brown, A., Camarata, J., Campanio, A., Chand, J., Chazaro, R., Choquet, Y., Colabelli, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeWitt, K., Dewar, K., Diaz, J., Dodge, S., Fero, S., Ferreira, P., Fitzgerald, W., Jace, D., Galagan, J., Gargya, S., Gierke, S., Geris, G., Giville, M., Graham, I., Grand Pierre, N., Hados, R., Harford, A., Horton, L., Hulse, W., Iliev, I., Johnson, R., Jones, C., Kumar, A., Karas, A., Kells, C., Lakorke, K., Lamazeres, K., Sanders, I., Schorsky, J., Levine, R., Liu, G., Maclean, M., McDonald, P., Major, J., Margolis, N., Matthews, C., McCarthy, M., McMan, P., McKenna, K., Moeckels, R., Melidj, J.,

Menais, L., Mhova, T., Menda, V., Murphy, T., Nasir, N., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, G., Peterson, K., Pinckard, P., Pierre, N., Pollard, V.,
 Raymond, C., Rella, R., Rieback, M., Riley, P., Risse, C., Rodov, P.,
 Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stancu-Thomant, N., Stojanovic, N.,
 Strauss, A., Subramanian, A., Talamas, J., Tostoy, S., Theodore, J.,
 Topham, K., Travers, M., Travers, N., Trillido, J., Vassiliou, H.,
 Viel, R., Vo, A., Wilson, R., Wu, X., Wyman, B., Ye, W., Yoon, J.,
 Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (32-Nov-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 190309)
 Birren, B., Nishium, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Birtu, N., Bastien, V., Bloom, T., Bonuslavsky, L., Boukharin, B.,
 Camarata, J., Chang, J., Chazarov, B., Cheng, Y., Collins, A.,
 Cook, A., Cooke, P., Deatellano, K., Iwar, K., Diaz, S., Jagan, S.,
 Piro, S., Ferreira, P., Fitzgerald, M., Gargano, T., Gargano, J.,
 Garbisa, S., Gird, S., Graham, L., Grant, P., Green, N., Hatz, N.,
 Hauss, B., Horton, L., Hume, W., Jilley, L., Johnson, R., Jones, C.,
 Kamei, A., Karakas, A., Kells, C., Landers, T., Levine, R.,
 Litchfield, R., Liu, G., Maclean, C., McDonald, P., Major, T.,
 Matthews, C., McCarthy, M., Melittim, J., Menais, L., Mhova, T.,
 Menda, V., Murphy, T., Naylor, J., Nguyen, C., Nishium, C., Norman, C.,
 Norbu, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, G.,
 Peters, G.K., Pinckard, P., Pierre, N., Raymond, C., Rella, R.,
 Risse, C., Rodov, P., Roman, J., Roy, A., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stancu-Thomant, N.,
 Stojanovic, N., Talamas, J., Tostoy, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliou, H., Viel, R., Vo, A., Wilson, R., Wu, X.,
 Wyman, B., Yoon, J., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (21-Oct-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 21, 2002 this sequence version replaced g1:2461725.
 All repeats were identified using RepeatMasker:
 Smith, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/MW/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WMR
 Web site: <http://www.seq.wi.mit.edu>
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 11418
 Center clone name: 491_LH15
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 189671 bases at least Q40
 Consensus quality: 189924 bases at least Q40
 Consensus quality: 189924 bases at least Q20
 Insert size: 172000; aarose+rp
 Insert size: 190099; sum-of-contigs
 quality coverage: 15.5 to Q20 bases; aarose+rp
 quality coverage: 14.0 to Q20 bases; sum-of-contigs

 * NOTE: This is a "working draft" sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 7203: contig G; 7204 bp in length
 * 7204 7303: gap of 100 bp
 * 7304 37157: contig G; 29854 bp in length
 * 37154 37257: gap of 100 bp
 * 37254 82905: contig G; 45449 bp in length
 * 82907 83005: gap of 98 bp

FEATURES * 83007 190309: contig of 107303 bp in length.
Location/Qualifiers
source 1. .190309
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-491H15"
/clone_lib="RPCI-24 Male Mouse BAC"
1. .7203
misc_feature /note="assembly_fragment"
7304. 37157
misc_feature /note="assembly_fragment"
37258. 82906
misc_feature /note="assembly_fragment"
83007. .190309
misc_feature /note="assembly_fragment"
BASE COUNT 57985 a 36628 c 36057 g 59339 t 300 others
ORIGIN
Query Match 53.3%; Score 27.2; DB 2; Length 190309;
Best Local Similarity 72.9%; Pred. No. 6;
Matches 35; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
OY 3 CTCCTTTAAGGCGTGAATGCTGCACTGTCATGTCACCTTAA 50
Db 120177 CTCCTTCAGGACTTGAAAGTCCATCTGCCAGGTGCACATTTCTAA 120130
RESULT 12
AC074174 183036 bp DNA linear HTG 15-JUL-2000
LOCUS AC074174
DEFINITION Mus musculus chromosome 5 clone RP23-123M9 strain C57BL6/J, WORKING
DRAFT SEQUENCE. 21 unordered pieces.
AC074174
AC074174.1 GI:9211284
VERSION HTG: HTGS_PHASE1: HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 183036)
Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,O.L., Maduro,V.B.,
Mastrian,S.D., McCloskey,J.C., Ojodu,M.A., Pearson,R.,
Stantirlop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
Wetherby,K.D. and Green,E.D.
NISC Mouse Sequencing Initiative
TITLE NISC Mouse Sequencing Initiative
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 183036)
Green,E.D.
COMMENT Direct Submission
Submitted (15-JUL-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouseseqhgrl.nih.gov
----- Project Information
Center project name: wx
Center clone name: 123M9
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168781 bases at least Q40
Consensus quality: 172931 bases at least Q30
Consensus quality: 175240 bases at least Q20
Insert size: 20000; agarose-gel
Insert size: 20200; pulse-field-gel
Insert size: 181036; sum-of-contigs

Quality coverage: 4.16x in Q20 bases; agarose-gel
Quality coverage: 4.12x in Q20 bases; pulse-field-gel
Quality coverage: 4.60x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
3005: contig of 3005 bp in length
3105: gap of unknown length
3106: 3105: gap of unknown length
5607: contig of 2502 bp in length
5608: 5707: gap of unknown length
5708: 8502: contig of 2795 bp in length
8503: 8602: gap of unknown length
8603: 11907: contig of 3305 bp in length
11908: 12007: gap of unknown length
12008: 14665: contig of 2658 bp in length
14666: 14765: gap of unknown length
14766: 17904: contig of 3139 bp in length
17904: 18004: gap of unknown length
18005: 22549: contig of 4545 bp in length
22550: 22649: gap of unknown length
22650: 27185: contig of 4536 bp in length
27186: 27285: gap of unknown length
31989: 31998: contig of 4713 bp in length
31999: 32098: gap of unknown length
32099: 36458: contig of 4358 bp in length
36459: 36556: gap of unknown length
36557: 41219: contig of 4663 bp in length
41220: 41319: gap of unknown length
41320: 45618: contig of 4299 bp in length
45619: 45718: gap of unknown length
45719: 53120: contig of 7402 bp in length
53121: 53220: gap of unknown length
53221: 59275: contig of 6055 bp in length
59276: 59375: gap of unknown length
59376: 65440: contig of 6065 bp in length
65441: 65540: gap of unknown length
65541: 76498: contig of 10958 bp in length
76499: 76598: gap of unknown length
76599: 92957: contig of 16359 bp in length
92958: 93057: gap of unknown length
93058: 109556: contig of 16499 bp in length
109557: 109656: gap of unknown length
109657: 128730: contig of 20074 bp in length
128731: 129830: gap of unknown length
129831: 147253: contig of 17423 bp in length
147254: 147353: gap of unknown length
147354: 183036: contig of 35683 bp in length.
Location/Qualifiers
1. .183036
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-123M9"
/clone_lib="RPCI mouse BAC library 23"
1. .3005
misc_feature /note="assembly_fragment"
3106. .5607
misc_feature /note="assembly_fragment"
5708. 8502
misc_feature /note="assembly_fragment"
8603. .11907
misc_feature /note="assembly_fragment"
12008. .14665
misc_feature /note="assembly_fragment"
14766. .17904

[illegible][illegible]

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
3 (bases 1 to 183291)
Worley, K.C.
Direct Submission
Submitted (11-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 183291)

REFERENCE
AUTHORS

Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Banks, T., Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jachson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, W., Louised, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Nickerson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenkwo, S., Ogih, M., Okwunu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williams, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
3 (bases 1 to 183291)
Worley, K.C.
Direct Submission
Submitted (11-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 183291)

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
3 (bases 1 to 183291)
Worley, K.C.
Direct Submission
Submitted (11-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 183291)

REFERENCE
AUTHORS

Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Banks, T., Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jachson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, W., Louised, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Nickerson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenkwo, S., Ogih, M., Okwunu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williams, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
3 (bases 1 to 183291)
Worley, K.C.
Direct Submission
Submitted (11-SEP-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 183291)

REFERENCE
AUTHORS

Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Banks, T., Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buha, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsl, F., Howard, S., Huber, J., Huliy, S., Hume, J., Jachson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, W., Louised, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Nickerson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenkwo, S., Ogih, M., Okwunu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williams, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

Query Match 51.0%; Score 26; DB 3; Length 183291;
Best Local Similarity 70.0%; Pred. NO. 18;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

```
misc_feature      245022..246206
BASE COUNT      62352 a 43072 c 44760 g 69919 t 35652 others
ORIGIN
Query Match      51.0%; Score 26; DB 2; Length 255755;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 35; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY      1 TGGTCCTTTTAAGGGCTTGAAATGCTGCAACTGTGATGTGTACACTTAA 50
      ||||| ||||| | ||||| || ||||| ||||| ||||| |||||
Db      190274 TGGTCCTTTTTCCTCCCGAAGCCTAGAACTTTCAGGTGTACTTTAA 190323
```

Search completed: August 20, 2003, 12:34:34
Job time : 267.82 secs

JOURNAL
TITLE Saurin, W., Bernot, A. and Weissenbach, J.
COMMENT Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
REFERENCE 10899143
GENOSCOPE 3 (bases 1 to 902)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr)

COMMENT
 Web : www.genoscope.cns.fr
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
SOURCE location/Qualifiers
 1. 902
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="037P04"
 /clone_1lb="G"
 /note="Genoscope sequence ID : CDBG037DH02SP1-end : POC-0r1"

BASE COUNT 127 a 156 c 300 g 92 t 227 others

Query Match 13.3% Score 50; DB 29; Length 902;
Best Local Similarity 31.6% Pred. No. 0.0053;
Matches 99; Conservative 67; Mismatches 147; Indels 0; Gaps 0;

4 CCCCCGACCCATCTGCTCCGCCCTTCTCTCAACCTCCCGAGACCTTAAGG 63
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 881 CCCCCCTCTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCSSGGGG 822
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 64 GCGCGCGGCGGCGAAGCGAGGCGCTGCGCTGACCCGAGGAAAGGCGCCAGTCTA 123
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 821 GCGGCGGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 762
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 124 GGTCTTAATGCGGTGCGCTCTCTTGACAGCGCGGCTTGGGGAACAAGCGGGAGC 183
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 761 CGSSGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 702
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 184 AGAGATAAGTGACATACAGAGAGATTTGTCGCGCGCTGATCTCTCCGACA 243
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 701 MMGGGCGCGMGGMVCMCTCKMKKTTKKKKKTKTKMKTKTKMKTKMKTKMKMK 642
 :
 244 GGAAGCGCGAGCTATTAAAGACCTATGATTTACTTTCTTCTGGAAGCTTCT 303
 :
 641 KKKMKGMKMKMATDATTCKDMAVATMAATAAKAADAAMATAKKAATKGA 582
 :
 304 TGGGAGAGACAA 316
 :
 581 AKTGTAKAAAMA 569

RESULT 2
LOCUS CD329933 982 bp mRNA linear EST 28-MAY-2003
DEFINITION AGENCOURT_14164463 NICHG_XGC_Brn1 Xenopus laevis cDNA clone
IMAGE 6951625 5', mRNA sequence.
ACCESSION CD329933
VERSION CD329933
KEYWORDS EST
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 982)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT Unpublished
CONTACT Daniela S. Gerhard, Ph.D.
OFFICE Office of Cancer Genomics
NATIONAL National Cancer Institute / NIH
BUILDING Bldg. 31 Rm10A07 Bethesda, MD 20892
EMAIL Email: cgapbs-tr@mail.nih.gov
TISSUE Tissue Procurement:
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM14570 row: p column: 24
 High quality sequence stop: 297.

FEATURES
SOURCE location/Qualifiers
 1. 982
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:6951625"
 /dev_stage="adult"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NICHG_XGC_Brn1"
 /note="Organ: brain; Vector: pCMV-SPORT6, Site: 1; NotI, Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT 172 a 221 c 445 g 89 t 55 others

Query Match 13.2% Score 49.4; DB 14; Length 982;
Best Local Similarity 43.2% Pred. No. 0.0079;
Matches 95; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

4 CCCCCGACCCCTATCTGCTCCGCCCTTCTCTCAACCTCCCGAGACCTTAAGG 63
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 543 CCCCCCGCCCCCCCCCCCCCCCCCTCCCTCCCGCTCCCGCCCCCCCCCGGG 484
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 64 GCGCGCGGCGGCGAAGCGAGGCGCTGCGCTGACCCGAGGAAAGGCGCCAGTCTA 123
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 483 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 424
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 124 GGTCTTAATGCGGTGCGCTCTCTTGACAGCGCGGCTTGGGGAACAAGCGGGAGC 183
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 423 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 364
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 184 AGAGATAAGTGACATACAGAGAGATTTGCTGCGCGG 223
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 363 NGNGNGNGNGNGNGNGNGNGNGNGNGNGNGNGNGNGNG 324

RESULT 3
LOCUS BX391246/c 1061 bp mRNA linear EST 13-MAY-2003
DEFINITION BX391246 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1008Y19 3-PRIME, mRNA sequence.
ACCESSION BX391246
VERSION BX391246
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1061)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
CONTACT Genoscope
COMMENT Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr


```

OY 4 CCCCCGACCCCTCATCTGCTGCCGCCCTTCTCTCCACCTCCCGAGACCCCTAAG 63
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1009 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 950
OY 64 GCGGCGGGGGCCCAAGCCGAGCGCTGCGCTGACCCCGAGCGGAAGGCCCACTCTA 123
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 949 CCCTTKGIVVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMVMMV 890
OY 124 GGTCTTAATGCGGGTGGCTCTCTTTGACAGCGCGCTTTGGGACACAGCGGGGAGC 183
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 889 AAMVAVGVWGGGGVGRMAAAMVRYAAAAAGAGGGGSGAGAAAAVSSSCGRAR 830
OY 184 AGAGATAAGTGATGACATACAGACAGATTTGTGCGCGCGCTGATCTCTCCGACA 243
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 829 ARAAMAVGVSSSSSSSSCGMAAAGMGREGVSSSSCCCSVRCMDVAVACSCMCMCA 770
OY 244 GGAACGCC 252
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 769 GMARCCCB 761

RESULT 6
LOCUS BX338963 514 bp mRNA linear EST 02-MAY-2003
DEFINITION BX338963 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
ACCESSION clone CSOD1065YP08 5-PRIME, mRNA sequence.
VERSION BX338963
KEYWORDS BX338963.1 GI:30347036
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang, Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue Genoscope sequence ID : CSOD1065DH04QP1.
LOCATION/Qualifiers
FEATURES
source 1..514
/mol_type="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1065YP08"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 100 a 70 c 159 g 101 t 84 others
ORIGIN
Query Match 12.1% Score 45.4; DB 13; Length 514;
Best Local Similarity 28.0%; Pred. No. 0.081;
Matches 53; Conservative 62; Mismatches 74; Indels 0; Gaps 0;
OY 7 CCGGACCCCTCATCTGCTGCCGCCCTTCTCTCCACCTCCCGAGACCCCTAAGGCG 66
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 323 CCGYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 382
OY 67 GCGGCGCCCAAGCGGGGCTGCTGACCCCGAGCGGAAGGGGCGCCAGTCTAGT 126
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 383 SSSSGGGGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 442
OY 127 CCTAATGCGGGTGGCTCTCTTTGACAGCGCGCTTTGGGACACAGCGGGGAGCAG 186

```

```

Db 443 GGGSGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 502
OY 187 GATAAGTG 195
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 503 GTGGGGGG 511

RESULT 7
LOCUS CNS01690/c 1201 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence Sp6 end of BAC
BACN15009 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL106446
VERSION AL106446.1 GI:5621964
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with The European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk/. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
LOCATION/Qualifiers
FEATURES
source 1..1201
/mol_type="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN15009"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"
BASE COUNT 252 a 289 c 224 g 107 t 329 others
ORIGIN
Query Match 12.0% Score 45; DB 29; Length 1201;
Best Local Similarity 41.5%; Pred. No. 0.14;
Matches 56; Conservative 32; Mismatches 47; Indels 0; Gaps 0;
OY 13 CCGTCATGCTGGCTGCCGCCCTTCTCTCCACCTCCCGAGACCCCTAAGGCGGGG 72
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1138 CCTCTCKCCGCBKSGCGCTGTGCGCCCTCCSCCGGSGCCTGSSGCGCGCGGSG 1079
OY 73 GCGCAAGCGGAGGGGCTGCTGACCCCGAGCGGAAGGGGCGCCAGTCTAATTA 132
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1078 GCSBSHCYGGGKGCGSCGKSGCBGCTGKBKKGSKKCCBKBKBSCKGKYVTT 1019
OY 133 GCGGGTGGCTCTCC 147
    |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1018 GCKGKKRSTRTKCC 1004

RESULT 8
LOCUS AL559222 941 bp mRNA linear EST 31-MAY-2003
DEFINITION AL559222 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL559222
VERSION AL559222.2 GI:31283355
KEYWORDS EST.

```


[illegible][illegible]


```
Db 181 ACAGAGATTAAGTGCATACAGAGAGATTGGTGGCGCGCTGATCTCTCCG 240
QY 241 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 300
Db 241 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 300
QY 301 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 360
Db 301 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 360
QY 361 CTGCACAGGCGAGC 375
Db 361 CTGCACAGGCGAGC 375
```

```
RESULT 2
US-09-761-466-3
; Sequence 3, Application US/09761466
; Patent No. US20020022259A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Ike W.
; APPLICANT: Izumo, Seigo
; TITLE OF INVENTION: Cardiac-Cell Specific Enhancer Elements
; FILE REFERENCE: 01948/069002
; CURRENT APPLICATION NUMBER: US/09/761,466
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,419
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-761-466-3
```

```
Query Match 100.0%; Score 375; DB 9; Length 1072;
Best Local Similarity 100.0%; Pred. No. 1.8e-112;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGGCCCCCGGACCCATCTGCGTCCGCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 60
Db 1 AGGCCCCCGGACCCATCTGCGTCCGCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 60
QY 61 AGGCGCGGCGGCGCCCAAGCGAGGCGCTGCGCTGACCCCGAGCGGAAGGCGCCAGT 120
Db 61 AGGCGCGGCGGCGCCCAAGCGAGGCGCTGCGCTGACCCCGAGCGGAAGGCGCCAGT 120
QY 121 CTAGTCTCTAATGCGGCTGCGTCTCTTGAACAGCGGCGTTGGGGAACAAGCGGGG 180
Db 121 CTAGTCTCTAATGCGGCTGCGTCTCTTGAACAGCGGCGTTGGGGAACAAGCGGGG 180
QY 181 ACAGAGATTAAGTGCATACAGAGAGATTGGTGGCGCGCTGATCTCTCCG 240
Db 181 ACAGAGATTAAGTGCATACAGAGAGATTGGTGGCGCGCTGATCTCTCCG 240
QY 241 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 300
Db 241 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 300
QY 301 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 360
Db 301 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 360
QY 361 CTGCACAGGCGAGC 375
Db 361 CTGCACAGGCGAGC 375
```

RESULT 3
US-09-761-466-4

```
; Sequence 4, Application US/09761466
; Patent No. US20020022259A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Ike W.
; APPLICANT: Izumo, Seigo
; TITLE OF INVENTION: Cardiac-Cell Specific Enhancer Elements
; FILE REFERENCE: 01948/069002
; CURRENT APPLICATION NUMBER: US/09/761,466
; CURRENT FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,419
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7838
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(7838)
; OTHER INFORMATION: n = A,T,C or G
US-09-761-466-4
```

```
Query Match 100.0%; Score 375; DB 9; Length 7838;
Best Local Similarity 100.0%; Pred. No. 3.5e-112;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AGGCCCCCGGACCCATCTGCGTCCGCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 60
Db 3588 AGGCCCCCGGACCCATCTGCGTCCGCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 3647
QY 61 AGGCGCGGCGGCGCCCAAGCGAGGCGCTGCGCTGACCCCGAGCGGAAGGCGCCAGT 120
Db 3648 AGGCGCGGCGGCGCCCAAGCGAGGCGCTGCGCTGACCCCGAGCGGAAGGCGCCAGT 3707
QY 121 CTAGTCTCTAATGCGGCTGCGTCTCTTGAACAGCGGCGTTGGGGAACAAGCGGGG 180
Db 3708 CTAGTCTCTAATGCGGCTGCGTCTCTTGAACAGCGGCGTTGGGGAACAAGCGGGG 3767
QY 181 ACAGAGATTAAGTGCATACAGAGAGATTGGTGGCGCGCTGATCTCTCCG 240
Db 3768 ACAGAGATTAAGTGCATACAGAGAGATTGGTGGCGCGCTGATCTCTCCG 3827
QY 241 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 300
Db 3828 ACAGAAACGGGAGCTTTAAAGACCTATGATTACTTATCTTCTCGAAAGCT 3887
QY 301 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 360
Db 3888 TCTTGGGAGAGACAAAGATGTCTCCCTGCTTAAGACACAAAGCCACAAAGAGGT 3947
QY 361 CTGCACAGGCGAGC 375
Db 3948 CTGCACAGGCGAGC 3962
```

```
RESULT 4
US-10-237-496-101/C
; Sequence 101, Application US/10237496
; Publication No. US20030138896A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Zhang, Zemin
```



```
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079669
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 101
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-243-402-101
```

```
Query Match          9.8%; Score 36.6; DB 12; Length 1111;
Best Local Similarity 49.7%; Pred. No. 0.093;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
```

```
QY 53 ACCCCCTAAGGGGGGGGGGGGGCCCAAGCCGAGGGGCTGGGCGCTGACCCCGAGGGAAG 112
DB 466 AGCAGGAAAGGGCCCGAGCCTCGCACCCAGCTGGGCGCTGGACAGCTCCGAGGCCAGC 407
QY 113 GCCCCTAAGGCTCTAATGCGGGGTGCGTCTCTTGACAGCGCGGCTTTGGGGAAC 172
DB 406 AGCCAGTCCCGGCATATATCCCGAGCGGGGTGATGAGAGGGCAGGGTTGGGGAACG 347
QY 173 CAGCGGGGACGAGATAGGTGACATACAGACAGATTGGTGGCGGCGCTATACTC 232
DB 346 CCCCCGCAAAAGCATTTGGGCTGCGGCGAGCGATTGACTTCCCTCAGAGTCCCGC 287
QY 233 CTCTCC 239
DB 286 CTCTCGC 280
```

```
RESULT 11
US-10-243-431-101/c
; Sequence 101, Application US/10243431
; Publication No. US20030138903A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Baton, Dan
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C37
; CURRENT APPLICATION NUMBER: US/10/243,431
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
```

```
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 101
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-243-431-101
```

```
Query Match          9.8%; Score 36.6; DB 12; Length 1111;
Best Local Similarity 49.7%; Pred. No. 0.093;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;
```

```
QY 53 ACCCCCTAAGGGGGGGGGGGGGCCCAAGCCGAGGGGCTGGGCGCTGACCCCGAGGGAAG 112
DB 466 AGCAGGAAAGGGCCCGAGCCTCGCACCCAGCTGGGCGCTGGACAGCTCCGAGGCCAGC 407
QY 113 GCCCCTAAGGCTCTAATGCGGGGTGCGTCTCTTGACAGCGGGGCTTTGGGGAAC 172
DB 406 AGCCAGTCCCGGCATATATCCCGAGCGGGGTGATGAGAGGGCAGGGTTGGGGAACG 347
QY 173 CAGCGGGGACGAGATAGGTGACATACAGACAGATTGGTGGCGGCGCTATACTC 232
DB 346 CCCCCGCAAAAGCATTTGGGCTGCGGCGAGCGATTGACTTCCCTCAGAGTCCCGC 287
QY 233 CTCTCC 239
DB 286 CTCTCGC 280
```

```
RESULT 12
US-10-245-164-101/c
; Sequence 101, Application US/10245164
; Publication No. US20030138904A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Baton, Dan
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C76
; CURRENT APPLICATION NUMBER: US/10/245,164
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
```



```
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 101
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-107-101

Query Match
Best Local Similarity 49.7%; Score 36.6; DB 14; Length 1111;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 53 ACCCTAAGGGGCGGGGCGCCCAAGCCGAGGGCGCTGACCCCGAGCGAAGG 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 466 ACCACGGAAGGGCCCGAGCCCTGCGACCCAGCTGGGCCCTGGCGAGCGTCCGACGCCAGC 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 GCCCAGTCTAGGTCTTATGCGGGTGCGCTCTCTTTGACAGCGCGCTTTGGGACAA 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 406 AGCCCACTCCCGCCATATCCCGGAGCGGGGTCTATGAGAGAGGCGAGGGTTGGGAAAG 347
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 CAGCGGGGACGAGATTAAGTGACATACAGACGAGATTGGTGGCGCGCTGATATCTC 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 346 CCCCCGCAAAAGCCATTGGCGGTACCTGCGGAGCGAGATTGACTTCCCTAGAGTCCCGC 287
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 CTCTCCC 239
   | | | | |
Db 286 CTCTCGC 280

RESULT 15
US-10-245-143-101/c
; Sequence 101, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
```

```
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 101
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-143-101

Query Match
Best Local Similarity 49.7%; Score 36.6; DB 14; Length 1111;
Matches 93; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 53 ACCCTAAGGGGCGGGGCGCCCAAGCCGAGGGCGCTGACCCCGAGCGAAGG 112
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 466 AGCAGGAAAGGGCCCGAGCCCTGCGACCCAGCTGGGCCCTGGCGAGCGTCCGAGGCCAGC 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 113 GCCCAGTCTAGGTCTTATGCGGGTGCGCTCTCTTTGACAGCGCGCTTTGGGACAA 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 406 AGCCCACTCCCGCCATATCCCGGAGCGGGGTCTATGAGAGAGGCGAGGGTTGGGAAAG 347
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 CAGCGGGGACGAGATTAAGTGACATACAGACGAGATTGGTGGCGCGCTGATATCTC 232
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 346 CCCCCGCAAAAGCCATTGGCGGTACCTGCGGAGCGAGATTGACTTCCCTAGAGTCCCGC 287
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 233 CTCTCCC 239
   | | | | |
Db 286 CTCTCGC 280
```

Search completed: August 20, 2003, 18:13:15
Job time : 348.431 secs


```
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 849
LENGTH: 1281
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (416)..(1033)
US-09-620-312D-849
```

Query Match 9.3%: Score 35; DB 4; Length 1281;

Best Local Similarity 56.5%: Pred. No. 0.32; 50; Indels 0; Gaps 0;

Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```
QY 5 CCCCCGACCCCTGATCTGCTCCGCCCTTCTCTCCACCCCTCCGACCCCTAAAGG 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 CTTCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCAAGC 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 GCGCGGGGGGCGGAGCGGCGGCTGCGCTGACCCCGGAGGAGGCGCCCG 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 GGGGCGGCGGCGGAGCGGCGGCGGCTGCGCTGCGGAGGAGGCGGCGCGG 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 3

```
US-09-819-989-3
Sequence 3, Application US/09819989
Patent No. 6482629
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: C1001200
CURRENT APPLICATION NUMBER: US/09/819,989
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 19650
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(19650)
OTHER INFORMATION: n = A,T,C or G
US-09-819-989-3
```

Query Match 9.0%: Score 33.6; DB 4; Length 19650;

Best Local Similarity 52.1%: Pred. No. 2.8;

Matches 75; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

```
QY 1 AGGCCCCCGACCCCTGATCTGCTCCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1885 AGAGCCCTGGGCGGCTCTCTCCCTTCCCTCCACCCCTCCCGGCGGCGGCTCCGC 1944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 AGGCGGCGGCGGCGGCGGCGGAGCGGCGGCTGCGCTTACCCCGAGCGGAGGCCCAAGT 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1945 GTGGGCGGCTGATCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2004
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 CTAGCTCTAATGGGCGGCTGCTC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2005 GCGGCGCTGATGCTGTGATGACGGC 2028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 4

```
US-09-620-312D-358
Sequence 358, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 358
LENGTH: 2780
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (219)..(2045)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(2780)
OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-358
```

Query Match 8.9%: Score 33.2; DB 4; Length 2780;

Best Local Similarity 53.0%: Pred. No. 1.6;

Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```
QY 52 GACCCCTAAGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 106 GCGGCGCTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 112 GCGCCCATGCTAGTCTCAATGCGGCGGCTGCTCTTGAACGCGGCGGCTTGGAGACA 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 ATCCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 172 ACAGCGGGAGAGAG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-5600
TELEFAX: 703-836-5288
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Caulobacter crescentus
STRAIN: CB 15
US-08-194-290-6

Query Match
Best Local Similarity 51.7%; Pred. No. 2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 43 ACCCTCCCGAGCCCTTAAGGGGGGGGGCCCAAGCCGAGGAGCGCTGCGCTGACCCC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1115 ACCGCCCTGAACACCAACACAGCGGCGGCTCAACCGTCACCGCGGCGTGGCCAG 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 GAGCGGAAGGGCCCGCAGTCTAGTCTTAATGCGGGTGGCGTCTCTTTGACAGCGGCGT 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1175 AACCTGACCGCCACAGACCGCGCTCAAGCGCGAACAACGTCGCGCGGCGGCC 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 TTGGGGACACAGCGGAGAGAG 187
DB 1235 AACGTACCGTCGCTCGACGCGCG 1259

RESULT 8
US-08-614-377A-6
; Sequence 6, Application US/08614377A
; Patent No. 5976864
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H.
; APPLICANT: No. 5976864ellini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF
; TITLE OF INVENTION: HETEROLOGOUS
; NUMBER OF INVENTION: POLYPEPTIDES FROM CAULOBACTER
; CORRESPONDENCE ADDRESS: 12
; ADDRESSEE: Fish & Richardson PC
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/614,377A
; FILING DATE: 12-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/194,290
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 435
; APPLICATION NUMBER: US 07/895,367
; FILING DATE: 09-JUNE-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsao, Y. Rocky
; REGISTRATION NUMBER: 34053
; REFERENCE/DOCKET NUMBER: 08106/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
```

```
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3300 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Caulobacter crescentus
STRAIN: CB 15
US-08-614-377A-6

Query Match
Best Local Similarity 51.7%; Pred. No. 2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 43 ACCCTCCCGAGCCCTTAAGGGGGGGGGCCCAAGCCGAGGAGCGCTGCGCTGACCCC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1115 ACCGCCCTGAACACCAACACAGCGGCGGCTCAACCGTCACCGCGGCGTGGCCAG 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 GAGCGGAAGGGCCCGCAGTCTAGTCTTAATGCGGGTGGCGTCTCTTTGACAGCGGCGT 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1175 AACCTGACCGCCACAGACCGCGCTCAAGCGCGAACAACGTCGCGCTGAGAGGGCGGCC 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 TTGGGGACACAGCGGAGAGAG 187
DB 1235 AACGTACCGTCGCTCGACGCGCG 1259
```

```
RESULT 9
US-09-142-648B-6
; Sequence 6, Application US/09142648B
; Patent No. 6210948
; GENERAL INFORMATION:
; APPLICANT: Smit, John
; APPLICANT: Bingle, Wade H.
; APPLICANT: No. 6210948ellini, John F.
; TITLE OF INVENTION: EXPRESSION AND SECRETION OF HETEROLOGOUS POLYPEPTIDES FROM CA
; FILE REFERENCE: 08106/002002
; CURRENT APPLICATION NUMBER: US/09/142,648B
; CURRENT FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: PCT/CA97/00167
; PRIOR FILING DATE: 1997-03-10
; PRIOR APPLICATION NUMBER: US 07/614,377
; PRIOR FILING DATE: 1996-03-12
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Caulobacter crescentus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)...(3178)
US-09-142-648B-6

Query Match
Best Local Similarity 51.7%; Pred. No. 2;
Matches 75; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 43 ACCCTCCCGAGCCCTTAAGGGGGGGGGCCCAAGCGGAGGAGCGCTGCGCTGACCCC 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1115 ACCGCCCTGAACACCAACACAGCGGCGGCTCAACCGTCACCGCGGCGTGGCCAG 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103 GAGCGGAAGGGCCCGCAGTCTAGTCTTAATGCGGGTGGCGTCTCTTTGACAGCGGCGT 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1175 AACCTGACCGCCACAGACCGCGCTCAAGCGCGAACAACGTCGCGCTGAGAGGGCGGCC 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 TTGGGGACACAGCGGAGAGAG 187
DB 1235 AACGTACCGTCGCTCGACGCGCG 1259
```



```

RESULT 10
US-09-152-060-38
SEQUENCE 40: APPLICATION 75/07152-00
PATENT NO. 6446245
GENERAL INFORMATION:
APPLICANT: ROSEN, G. H.
TITLE OF INVENTION: 2d Human Secreted Proteins
FILE REFERENCE: P2004105
CURRENT APPLICATION NUMBER: 08/37152-00
EARLIER FILING DATE: 1998-09-11
EARLIER APPLICATION NUMBER: 08/0584/0488
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/040752
EARLIER FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 60/040710
EARLIER FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 60/050934
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048109
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048357
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/048189
EARLIER FILING DATE: 1997-05-30
EARLIER APPLICATION NUMBER: 60/057765
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/048370
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/038358
EARLIER FILING DATE: 1997-12-14
NUMBER OF SEQ ID NOS: 118
SEQUENCE 1: 1-118
SEQ ID NO 68
LENGTH: 1762
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (6%)
OTHER INFORMATION: Duplicates 4,7,4, or 0
US-09-152-060-38

Query Match          87% Score 32.8; DB 4; Length 1762;
Best Local Similarity 55.2%; Pred. No. 17;
Matches 64; Conservative 0; Mismatches 52; Indels 0; Caps 0;

QY 457 TATPAAAAAATACTGATATTATTTTCGTGGTAAAGCTCTGGCGACACCA 315
DB 1440 TACTTAATATCTGATATATATATCTGCTTATGTGAATGATCATATGGGAAACAT 1484
YY 417 AATATTTCTTGTTATTAAGAATACAAATGATGATGAGGAAATTTCTGTAAAGGTA 372
DB 1490 GAATATATCTGATTAAGATATGATTAAGATATGATGATGATGATGATGATGAT 1545

RESULT 11
US-09-252-941A-1452
SEQUENCE 1452: APPLICATION 08/0242422A
PATENT NO. 6551795
GENERAL INFORMATION:
APPLICANT: MARRAS, JAMES E.
TITLE OF INVENTION: NATURAL AND AMINO ACID SEQUENCES RELATING TO INSULIN ANTAGONISTS
FILE REFERENCE: 107104140
CURRENT APPLICATION NUMBER: 08/0242422A
EARLIER FILING DATE: 1999-02-14
EARLIER APPLICATION NUMBER: US 60/374748
PRIOR FILING DATE: 1999-02-14
PRIOR APPLICATION NUMBER: 08/5744114
EARLIER FILING DATE: 1999-07-12
NUMBER OF SEQ ID NOS: 43142
SEQ ID NO 14552

```

[illegible]

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,434A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weseman, James C.
REGISTRATION NUMBER: 30,507
REFERENCE/DOCKET NUMBER: P00060US0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-3604
TELEFAX: 619-236-1048
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2875 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-458-434A-4

Query Match 8.6%; Score 32.4; DB 3; Length 2875;
Best Local Similarity 50.6%; Pred. No. 2.9; Indels 0; Gaps 0;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 5 CCCCCGACCTTCATCTGCTCCGCCCTTCTCTCCACCTCCGCGACCCCTAAAGG 64
DB 1791 CCCCCGACCTTCATCTGCTCCGCCCTTCTCTCCACCTCCGCGACCCCTAAAGG 1850
QY 65 GCGCGGGGCGCCCAAGCGAGGGGCGCTGCGCTGACCCCGGAGGGGCGCCAGTCTAG 124
DB 1851 CCGCGGGGCGGAGTCCCGCGCGCTGCTCTCCAGGGGAGGTGTTCGCCACGCGCGGA 1910
QY 125 GTCTTAATGCGGGTGGCGTCTCTTTGACAGGCG 158
DB 1911 GGGAGCGCGCAGCGCGGCTCTCTTTAAAGCGC 1944

RESULT 14
US-08-458-434A-6
Sequence 6, Application US/08458434A
Patent No. 6083690
GENERAL INFORMATION:
APPLICANT: Harris Ph.D., Stephen E.
APPLICANT: Mundy M.D., Gregory R.
APPLICANT: Gosh-Choudhury Ph.D., Nandini
APPLICANT: Feng Ph.D., Jian Q.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING
TITLE OF INVENTION: OSTEOGENIC AGENTS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: James C. Weseman, Esq.
STREET: 401 B. Street, Suite 1700
CITY: San Diego
STATE: CA
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,434A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weseman, James C.
REGISTRATION NUMBER: 30,507
REFERENCE/DOCKET NUMBER: P00060US0

TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-3604
TELEFAX: 619-236-1048
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-458-434A-6

Query Match 8.6%; Score 32.4; DB 3; Length 15144;
Best Local Similarity 50.6%; Pred. No. 5.9;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 5 CCCCCGACCTTCATCTGCTCCGCCCTTCTCTCCACCTCCGCGACCCCTAAAGG 64
DB 1791 CCCCCGACCTTCATCTGCTCCGCCCTTCTCTCCACCTCCGCGACCCCTAAAGG 1850
QY 65 GCGCGGGGCGCCCAAGCGAGGGGCGCTGCGCTGACCCCGGAGGGGCGCCAGTCTAG 124
DB 1851 CCGCGGGGCGGAGTCCCGCGCGCTGCTCTCCAGGGGAGGTGTTCGCCACGCGCGGA 1910
QY 125 GTCTTAATGCGGGTGGCGTCTCTTTGACAGGCG 158
DB 1911 GGGAGCGCGCAGCGCGGCTCTCTTTAAAGCGC 1944

RESULT 15
US-09-252-991A-12955
Sequence 12955, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT FILING DATE: 1999-02-18
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-07-27
PRIOR APPLICATION NUMBER: US 60/094,190
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 12955
LENGTH: 846
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12955

Query Match 8.6%; Score 32.2; DB 4; Length 846;
Best Local Similarity 61.2%; Pred. No. 1.9;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 2 GGGCCCCGACCTTCATCTGCTCCGCCCTTCTCTCCACCTCCGAGCCCTAA 61
DB 14 GGACACCGGGCGCATGTTATTTCACAGCGCTTCTGCGCTTCGCGAGCCAGAAC 73
QY 62 GGGCGCGCGGGCGCCAGCGGAGG 86
DB 74 AGGTGGGGCGCGGAGCGGCGG 98

Search completed: August 20, 2003, 14:32:38
Job time: 53.7146 secs

[illegible][illegible]

XX The present sequence is that of the human ZFP36 (homologous to zfp-36 in
CC mouse) gene that encodes the ZFP36 protein of the invention (see #
CC AB080316). The specification describes a newly isolated polynucleotide
CC comprising a sequence which is a polymorphic variant (PV) of a reference
CC sequence for the ZFP36 gene or its fragment and its encoded protein. The
CC ZFP36 polynucleotides and polypeptides have antirheumatic,
CC immunosuppressive and antiarthritic activities. The ZFP36 polypeptide is
CC useful for screening drugs targeting the ZFP36 polypeptide. ZFP36
CC isogenes or haplotype pairs are useful for improving the efficiency and
CC reliability of the discovery and development of drugs for treating
CC diseases associated with ZFP36 activity, e.g., autoimmune diseases such
CC as rheumatoid arthritis. Haplotyping the ZFP36 gene in an individual
CC gives useful information for validating ZFP36 as a candidate target for
CC treating a specific condition predicted to be associated with ZFP36
CC activity. Genotyping the ZFP36 gene of an individual can give information
CC used for developing diagnostic tests and therapeutic treatments. The
CC isolated polynucleotide is useful in studying the expression and function
CC of ZFP36 and in drug screening. Antibodies specific for the ZFP36 protein
CC are useful in many diagnostic and prognostic formats and therapeutic
CC methods. A recombinant non-human organism transformed with the ZFP36 gene
CC is useful in studying expression of the ZFP36 isogenes in vivo, for in
CC vivo drug screening and testing. Allele-specific oligonucleotides (ASO)
CC are useful as probes and primers and for assaying a polymorphism in the
CC target region.

XX
SQ Sequence 3889 BP; 648 A; 1304 C; 1078 G; 850 T; 9 other;

Query Match 11.7%; Score 43.8; DB 24; Length 3889;
Best Local Similarity 59.1%; Pred. No. 0.012;
Matches 75; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

OY 2 GGCCCCCGACCTCATCTGGCTCCGCCCTTCTCCACCTCCGAGCCCTAA 61
DB 301 GGCGAGCGCTCCCGCATCCGACACCCACCTTCCCGACGATTCCTCGGTGTCAC 360
OY 62 GGGCGCGGGGCGCCAGCGGCGCTCGGCTGACCCCGAGCGGAGGCCCGCAGTC 121
DB 361 GGCTGTCCACCGGCGCAAGCTCAGGCGGCTCTCCAGCGCGGCGGCGAAGGACCGAGTC 420

OY 122 TAGGTCC 128
DB 421 CAGGGCC 427

RESULT 5
ABK91535
ID ABR91535 standard; DNA; 7552 BP.

XX
AC ABR91535;

XX
DT 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #87.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KM acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX
OS Synthetic.

XX
PN WO200232943-A2.

XX
PD 25-APR-2002.

XX
PF 14-AUG-2001; 2001WO-US25721.

XX
PR 14-AUG-2000; 2000US-225097P.

XX
PR 14-NOV-2000; 2000US-252115P.

XX
PR 28-MAR-2001; 2001US-279257P.

XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
(CHAD/) CHADRABARTI B K.

XX
PI Nabel GJ, Huang Y;

XX
DR WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -

XX
PS Disclosure; page 577-579; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABR91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 7552 BP; 1797 A; 2149 C; 1995 G; 1611 T; 0 other;

Query Match 10.1%; Score 38; DB 24; Length 7552;
Best Local Similarity 62.8%; Pred. No. 0.78;
Matches 59; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

OY 26 TCCGCGCCCTTCTCTCCACCTCCCGAGACCCCTAAAGGGGGGGGCCCGACCGAGG 85
DB 4066 TCCGAGCCCTGCTCCCGCGCGAGCCGCGCCCGCCCGCCCGAGGACCGAGAGAGG 4125
OY 86 GCGCTGCGCTTACCCTCCCGAGCGGAGGCGCCAG 119
DB 4126 GCGGCGAGCGGACCGCGAGCCGCGAGCGGCGCCAG 4159

RESULT 6

ABK70011/c
ID ABR70011 standard; DNA; 1111 BP.

XX
AC ABR70011;

XX
DT 15-JUL-2002 (first entry)

DE cDNA encoding human Pro peptide #51.

XX Human; ss; gene; PRO; secreted protein; transmembrane protein;
KW genetic disorder; tumour; cancer.

XX
OS Homo sapiens.

XX
PN WO200224888-A2.

XX
PD 28-MAR-2002.

XX
PF 29-AUG-2001; 2001WO-US27099.

XX
PR 01-SEP-2000; 2000US-229896P.

XX
PR 05-SEP-2000; 2000US-230621P.

XX
PR 22-SEP-2000; 2000US-235147P.

XX
PR 10-NOV-2000; 2000WO-US30873.

XX
PR 12-JAN-2001; 2001US-261878P.

XX
PR 16-JAN-2001; 2001US-261910P.

XX
PR 16-JAN-2001; 2001US-262150P.

XX
PR 25-JAN-2001; 2001US-264395P.

XX
PR 02-FEB-2001; 2001US-266421P.

XX
PR 09-FEB-2001; 2001US-267623P.

XX
PR 28-FEB-2001; 2001WO-US06520.

XX
PR 09-MAR-2001; 2001US-274399P.

XX 23-JAN-2002 (first entry)
DT Human nervous system related polynucleotide SEQ ID NO 12346.
XX
DE
XX
XX Human; nocotropic; neuroprotective; cytostatic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnereary;
KW antiparkinsonian; antistickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX Homo sapiens.
OS
XX WO200159063-A2.
PN
XX
XX
PD 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
PF
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234423.
PR 21-SEP-2000; 2000US-0234424.
PR 21-SEP-2000; 2000US-0234957.
PR 25-SEP-2000; 2000US-0234958.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.

PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251988.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;

XX
 DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure: SEQ ID NO 24858; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC the nucleic acids may be used to produce the secreted (I), by inserting the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SO Sequence 28995 BP; 6151 A; 8327 C; 8051 G; 6466 T; 0 other;
 Query Match 9.4%; Score 35.2; DB 22; Length 28995;
 Best Local Similarity 58.7%; Pred. No. 8.6;
 Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
 QY 2 GGGCCCCGCGACCTGCTCCGCCCTTCTCTCCACCCCTCCGAGCCCTAA 61
 Db 4981 GACCTCCGCTCCCTCCCTCCCTCCCTCCGACATCTGAGTCCGAGCCATCTG 5040
 QY 62 GGGGCGGCGGCGCCCAAGCGAGGCGCTGCGCTGACCCGAG 105
 Db 5041 CTGAAAGAGGTGGCGCTGCGCTGCGCTGCGCTGCTGCGCG 5084
 RESULT 10
 AAK79967/C
 ID AAK79967 standard; DNA; 28995 BP.
 XX
 AC AAK79967;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34779.
 XX
 KW Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01354.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 07-MAY-2000; 2000US-0205515.
 PR 19-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.

REFERENCE 2 (bases 1 to 81512)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 81512)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 4 (bases 1 to 81512)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (01-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 5 (bases 1 to 81512)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (28-JAN-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Jan 28, 2003 this sequence version replaced gi:20376952.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.3% of Sequence.
 Estimated Total Number of Errors is 0.4.
 NOTE: This insert is not the entire sequence of the clone (entire sequence is 139.5b). It is clipped at the overlaps with AC008378 and AC106731. The number of bases overlapped with AC008378 is 4442 bps and with AC106731 is 15413 bps.

FEATURES
 source
 1. 81512
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CNC-281H14"

BASE COUNT 20011 a 19534 c 19980 g 21987 t

Query Match 96.5%; Score 362; DB 9; Length 81512;
 Best Local Similarity 99.2%; Pred. No. 1.3e-85;

Matches 375; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Y 1 AGGCCCCCGACCTCTATCTGCTCCGCGCTTCTCTCCACCTCCGAGCCCTAA 60
 D 54172 AGGCCCCCGACCTCTATCTGCTCCGCGCTTCTCTCCACCTCCGAGCCCTAA 54113
 Y 61 AGGGGCGGGGGGCGCCAAAGCCGAGGGGCGCTGCGCTGACCCCGAGCGAAAGGCCCACT 120
 D 54112 AGGGGCGGGGGGCGCCAAAGCCGAGGGGCGCTGCGCTGACCCCGAGCGAAAGGCCCACT 54053
 Y 121 CTAGTCTTAATGCGGGTGCCTCTTTGACAGCGGGCGCTTTGGGACAACAGCGGGG 180
 D 54052 CTAGTCTTAATGCGGGTGCCTCTTTGACAGCGGGCGCTTTGGGACAACAGCGGGG 53993
 Y 181 ACGAGAGTAAAGTGACATACACAGAGATTTGGTGGCGGCGGTGATCTCTCTCCG 240
 D 53992 ACGAGAGTAAAGTGACATACACAGAGATTTGGTGGCGGCGGTGATCTCTCTCCG 53933
 Y 241 ACAGGAAACGGGAGCTATTTAAAGACCTATGATTACTTTATCTTTCT--GAAA 297
 D 53932 ACAGGAAACGGGAGCTATTTAAAGACCTATGATTACTTTATCTTTCTCGAGAGAA 53873
 Y 298 GCTTCTTGGGAGAGACAAGAGATTTCCCTGCTTAAAGACACAGGCGACACAGCGAG 357
 D 53872 GCTTCTTGGGAGAGACAAGAGATTTCCCTGCTTAAAGACACAGGCGACACAGCGAG 53813
 Y 358 GGTCTGACAGGCGAGCG 375
 D 53812 GGTCTGACAGGCGAGCG 53795

RESULT 2
 LOCUS AL669924/c 137345 bp DNA linear PRI 24-APR-2002
 DEFINITION Human DNA sequence from clone XHbac-2L19 on chromosome 6, complete sequence.
 ACCESSION AL669924
 VERSION AL669924.9 GI:20068712
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 AUTHORS Peck, A.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk humquerry@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:19572797.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction primary digestion. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep_xhac-2L19_1s from a CHORI-501 human bac - PGF cell line library VECTOR: PTARBAC2.1
 This sequence was generated from part of bacterial clone contigs constructed by the MHC Haplotype Consortium and collaborators. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6/MHC.

FEATURES

source

1. 137345
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="XHbac-2L19"
 /clone_11b="CHORI-501"

BASE COUNT 33377 a 33207 c 33325 g 37436 t

ORIGIN

Query Match 96.5%; Score 362; DB 9; Length 137345;
 Best Local Similarity 99.2%; Pred. No. 1.4e-85;

Matches 375; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Y 1 AGGCCCCCGACCTCTATCTGCTCCGCGCTTCTCTCCACCTCCGAGCCCTAA 60
 D 54398 AGGCCCCCGACCTCTATCTGCTCCGCGCTTCTCTCCACCTCCGAGCCCTAA 54339
 Y 61 AGGGGCGGGGGGCGCCAAAGCCGAGGGGCGCTGCGCTGACCCCGAGCGAAAGGCCCACT 120
 D 54338 AGGGGCGGGGGGCGCCAAAGCCGAGGGGCGCTGCGCTGACCCCGAGCGAAAGGCCCACT 54279
 Y 121 CTAGTCTTAATGCGGGTGCCTCTTTGACAGCGGGCGTTTGGGACAACAGCGGGG 180
 D 54278 CTAGTCTTAATGCGGGTGCCTCTTTGACAGCGGGCGTTTGGGACAACAGCGGGG 54219
 Y 181 ACGAGAGTAAAGTGACATACACAGAGATTTGGTGGCGGCGGTGATCTCTCTCCG 240

[illegible]

1111.1	1111.2	1111.3	1111.4	1111.5	1111.6	1111.7	1111.8	1111.9	1111.10	1111.11	1111.12	1111.13	1111.14	1111.15	1111.16	1111.17	1111.18	1111.19	1111.20	1111.21	1111.22	1111.23	1111.24	1111.25	1111.26	1111.27	1111.28	1111.29	1111.30	1111.31	1111.32	1111.33	1111.34	1111.35	1111.36	1111.37	1111.38	1111.39	1111.40	1111.41	1111.42	1111.43	1111.44	1111.45	1111.46	1111.47	1111.48	1111.49	1111.50	1111.51	1111.52	1111.53	1111.54	1111.55	1111.56	1111.57	1111.58	1111.59	1111.60	1111.61	1111.62	1111.63	1111.64	1111.65	1111.66	1111.67	1111.68	1111.69	1111.70	1111.71	1111.72	1111.73	1111.74	1111.75	1111.76	1111.77	1111.78	1111.79	1111.80	1111.81	1111.82	1111.83	1111.84	1111.85	1111.86	1111.87	1111.88	1111.89	1111.90	1111.91	1111.92	1111.93	1111.94	1111.95	1111.96	1111.97	1111.98	1111.99	1112.00	1112.01	1112.02	1112.03	1112.04	1112.05	1112.06	1112.07	1112.08	1112.09	1112.10	1112.11	1112.12	1112.13	1112.14	1112.15	1112.16	1112.17	1112.18	1112.19	1112.20	1112.21	1112.22	1112.23	1112.24	1112.25	1112.26	1112.27	1112.28	1112.29	1112.30	1112.31	1112.32	1112.33	1112.34	1112.35	1112.36	1112.37	1112.38	1112.39	1112.40	1112.41	1112.42	1112.43	1112.44	1112.45	1112.46	1112.47	1112.48	1112.49	1112.50	1112.51	1112.52	1112.53	1112.54	1112.55	1112.56	1112.57	1112.58	1112.59	1112.60	1112.61	1112.62	1112.63	1112.64	1112.65	1112.66	1112.67	1112.68	1112.69	1112.70	1112.71	1112.72	1112.73	1112.74	1112.75	1112.76	1112.77	1112.78	1112.79	1112.80	1112.81	1112.82	1112.83	1112.84	1112.85	1112.86	1112.87	1112.88	1112.89	1112.90	1112.91	1112.92	1112.93	1112.94	1112.95	1112.96	1112.97	1112.98	1112.99	1113.00	1113.01	1113.02	1113.03	1113.04	1113.05	1113.06	1113.07	1113.08	1113.09	1113.10	1113.11	1113.12	1113.13	1113.14	1113.15	1113.16	1113.17	1113.18	1113.19	1113.20	1113.21	1113.22	1113.23	1113.24	1113.25	1113.26	1113.27	1113.28	1113.29	1113.30	1113.31	1113.32	1113.33	1113.34	1113.35	1113.36	1113.37	1113.38	1113.39	1113.40	1113.41	1113.42	1113.43	1113.44	1113.45	1113.46	1113.47	1113.48	1113.49	1113.50	1113.51	1113.52	1113.53	1113.54	1113.55	1113.56	1113.57	1113.58	1113.59	1113.60	1113.61	1113.62	1113.63	1113.64	1113.65	1113.66	1113.67	1113.68	1113.69	1113.70	1113.71	1113.72	1113.73	1113.74	1113.75	1113.76	1113.77	1113.78	1113.79	1113.80	1113.81	1113.82	1113.83	1113.84	1113.85	1113.86	1113.87	1113.88	1113.89	1113.90	1113.91	1113.92	1113.93	1113.94	1113.95	1113.96	1113.97	1113.98	1113.99	1114.00	1114.01	1114.02	1114.03	1114.04	1114.05	1114.06	1114.07	1114.08	1114.09	1114.10	1114.11	1114.12	1114.13	1114.14	1114.15	1114.16	1114.17	1114.18	1114.19	1114.20	1114.21	1114.22	1114.23	1114.24	1114.25	1114.26	1114.27	1114.28	1114.29	1114.30	1114.31	1114.32	1114.33	1114.34	1114.35	1114.36	1114.37	1114.38	1114.39	1114.40	1114.41	1114.42	1114.43	1114.44	1114.45	1114.46	1114.47	1114.48	1114.49	1114.50	1114.51	1114.52	1114.53	1114.54	1114.55	1114.56	1114.57	1114.58	1114.59	1114.60	1114.61	1114.62	1114.63	1114.64	1114.65	1114.66	1114.67	1114.68	1114.69	1114.70	1114.71	1114.72	1114.73	1114.74	1114.75	1114.76	1114.77	1114.78	1114.79	1114.80	1114.81	1114.82	1114.83	1114.84	1114.85	1114.86	1114.87	1114.88	1114.89	1114.90	1114.91	1114.92	1114.93	1114.94	1114.95	1114.96	1114.97	1114.98	1114.99	1115.00	1115.01	1115.02	1115.03	1115.04	1115.05	1115.06	1115.07	1115.08	1115.09	1115.10	1115.11	1115.12	1115.13	1115.14	1115.15	1115.16	1115.17	1115.18	1115.19	1115.20	1115.21	1115.22	1115.23	1115.24	1115.25	1115.26	1115.27	1115.28	1115.29	1115.30	1115.31	1115.32	1115.33	1115.34	1115.35	1115.36	1115.37	1115.38	1115.39	1115.40	1115.41	1115.42	1115.43	1115.44	1115.45	1115.46	1115.47	1115.48	1115.49	1115.50	1115.51	1115.52	1115.53	1115.54	1115.55	1115.56	1115.57	1115.58	1115.59	1115.60	1115.61	1115.62	1115.63	1115.64	1115.65	1115.66	1115.67	1115.68	1115.69	1115.70	1115.71	1115.72	1115.73	1115.74	1115.75	1115.76	1115.77	1115.78	1115.79	1115.80	1115.81	1115.82	1115.83	1115.84	1115.85	1115.86	1115.87	1115.88	1115.89	1115.90	1115.91	1115.92	1115.93	1115.94	1115.95	1115.96	1115.97	1115.98	1115.99	1116.00	1116.01	1116.02	1116.03	1116.04	1116.05	1116.06	1116.07	1116.08	1116.09	1116.10	1116.11	1116.12	1116.13	1116.14	1116.15	1116.16	1116.17	1116.18	1116.19	1116.20	1116.21	1116.22	1116.23	1116.24	1116.25	1116.26	1116.27	1116.28	1116.29	1116.30	1116.31	1116.32	1116.33	1116.34	1116.35	1116.36	1116.37	1116.38	1116.39	1116.40	1116.41	1116.42	1116.43	1116.44	1116.45	1116.46	1116.47	1116.48	1116.49	1116.50	1116.51	1116.52	1116.53	1116.54	1116.55	1116.56	1116.57	1116.58	1116.59	1116.60	1116.61	1116.62	1116.63	1116.64	1116.65	1116.66	1116.67	1116.68	1116.69	1116.70	1116.71	1116.72	1116.73	1116.74	1116.75	1116.76	1116.77	1116.78	1116.79	1116.80	1116.81	1116.82	1116.83	1116.84	1116.85	1116.86	1116.87	1116.88	1116.89	1116.90	1116.91	1116.92	1116.93	1116.94	1116.95	1116.96	1116.97	1116.98	1116.99	1117.00	1117.01	1117.02	1117.03	1117.04	1117.05	1117.06	1117.07	1117.08	1117.09	1117.10	1117.11	1117.12	1117.13	1117.14	1117.15	1117.16	1117.17	1117.18	1117.19	1117.20	1117.21	1117.22	1117.23	1117.24	1117.25	1117.26	1117.27	1117.28	1117.29	1117.30	1117.31	1117.32	1117.33	1117.34	1117.35	1117.36	1117.37	1117.38	1117.39	1117.40	1117.41	1117.42	1117.43	1117.44	1117.45	1117.46	1117.47	1117.48	1117.49	1117.50	1117.51	1117.52	1117.53	1117.54	1117.55	1117.56	1117.57	1117.58	1117.59	1117.60	1117.61	1117.62	1117.63	1117.64	1117.65	1117.66	1117.67	1117.68	1117.69	1117.70	1117.71	1117.72	1117.73	1117.74	1117.75	1117.76	1117.77	1117.78	1117.79	1117.80	1117.81	1117.82	1117.83	1117.84	1117.85	1117.86	1117.87	1117.88	1117.89	1117.90	1117.91	1117.92	1117.93	1117.94	1117.95	1117.96	1117.97	1117.98	1117.99	1118.00	1118.01	1118.02	1118.03	1118.04	1118.05	1118.06	1118.07	1118.08	1118.09	1118.10	1118.11	1118.12	1118.13	1118.14	1118.15	1118.16	1118.17	1118.18	1118.19	1118.20	1118.21	1118.22	1118.23	1118.24	1118.25	1118.26	1118.27	1118.28	1118.29	1118.30	1118.31	1118.32	1118.33	1118.34	1118.35	1118.36	1118.37	1118.38	1118.39	1118.40	1118.41	1118.42	1118.43	1118.44	1118.45	1118.46	1118.47	1118.48	1118.49	1118.50	1118.51	1118.52	1118.53	1118.54	1118.55	1118.56	1118.57	1118.58	1118.59	1118.60	1118.61	1118.62	1118.63	1118.64	1118.65	1118.66	1118.67	1118.68	1118.69	1118.70	1118.71	1118.72	1118.73	1118.74	1118.75	1118.76	1118.77	1118.78	1118.79	1118.80	1118.81	1118.82	1118.83	1118.84	1118.85	1118.86	1118.87	1118.88	1118.89	1118.90	1118.91	1118.92	1118.93	1118.94	1118.95	1118.96	1118.97	1118.98	1118.99	1119.00	1119.01	1119.02	1119.03	1119.04	1119.05	1119.06	1119.07	1119.08	1119.09	1119.10	1119.11	1119.12	1119.13	1119.14	1119.15	1119.16	1119.17	1119.18	1119.19	1119.20	1119.21	1119.22	1119.23	1119.24	1119.25	1119.26	1119.27	1119.28	1119.29	1119.30	1119.31	1119.32	1119.33	1119.34	1119.35	1119.36	1119.37	1119.38	1119.39	1119.40	1119.41	1119.42	1119.43	1119.44	1119.45	1119.46	1119.47	1119.48	1119.49	1119.50	1119.51	1119.52	1119.53	1119.54	1119.55	1119.56	1119.57	1119.58	1119.59	1119.60	1119.61	1119.62	1119.63	1119.64	1119.65	1119.66	1119.67	1119.68	1119.69	1119.70	1119.71	1119.72	1119.73	1119.74	1119.75	1119.76	1119.77	1119.78	1119.79	1119.80	1119.81	1119.82	1119.83	1119.84	1119.85	1119.86	1119.87	1119.88	1119.89	1119.90	1119.91	1119.92	1119.93	1119.94	1119.95	1119.96	1119.97	1119.98	1119.99	1120.00	1120.01	1120.02	1120.03	1120.04	1120.05	1120.06	1120.07	1120.08	1120.09	1120.10	1120.11	1120.12	1120.13	1120.14	1120.15	1120.16	1120.17	1120.18	1120.19	1120.20	1120.21	1120.22	1120.23	1120.24	1120.25	1120.26	1120.27	1120.28	1120.29	1120.30	1120.31	1120.32	1120.33	1120.34	1120.35	1120.36	1120.37	1120.38	1120.39	1120.40	1120.41	1120.42	1120.43	1120.44	1120.45	1120.46	1120.47	1120.48	1120.49	1120.50	1120.51	1120.52	112
--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	---------	-----

misc_feature	1..1023	/note="wgs_end_extension clone_end:Sp6"
misc_feature	2382..3456	/note="clone_boundary clone_end:Sp6 site: end_sequence:82214718"
misc_feature	118153..119195	/note="wgs_contig"
BASE COUNT	43967 a 40091 c 39195 g 41033 t 6824 others	
ORIGIN		
Query Match	61.5%, Score 230.8; DB 2; Length 171110;	
Best Local Similarity	89.7%; Pred. No. 1.5e-50;	
Matches 271;	Conservative 0; Mismatches 27; Indels 4; Gaps 2;	
QY	78 AGCCGAGGCGCGCTGCCTGACCCCGAGCGGAAAGGGCCCCATCTAGGCTCAATGCGGG 137	
Db	142378 AGCCGAGGCGCGCGCTGCCTGACCCCGAGCGGAAAGGGCCCCATCTAGGCTCAATGCGGG 142319	
QY	138 TGGCGTCCTCTTTGACAGCGCGCGCTTTGGGAGCAACAGCGGGAGAGATTAAGGTGAC 197	
Db	142318 TGGCGTCCTCTTTGACAGCGCGCGCTTTGGGAGCAACAGCGGGAGAGATTAAGGTGAC 142259	
QY	198 ATACCAGAGCAGATTTGGTGGCGCGCGCTGATATCTCTCTCCGACAGCAAGCGGAGGCT 257	
Db	142258 ATACCAGAGCAGATTTGGTGGCGCGCGCTGATATCTCTCTCCGACAGCAAGCGGAGGCT 142199	
QY	258 ATTTT-AAAGACCTATCGATTACTTATCTTTC--TGAAAGCTTCTTGGCGGAGAGA 313	
Db	142198 ATTTTAAAGACCTATCGATTACTTATCTTTC--TGAAAGCTTCTTGGCGGAGAGA 142139	
QY	314 CAAAGATGTTCCCTGCGCTAAAGACACAGCGCACACAGCGAGGGTGTGCACAGCGGAC 373	
Db	142138 CAAAGATGTTCCCTGCGCTAAAGATCAAGGCGCACACAGCGAGGGTGTGTACAGCGGAC 142079	
QY	374 GC 375	
Db	142078 GC 142077	
RESULT 4		
LOCUS	AC102990/c	241724 bp DNA linear HTG 10-MAY-2003
DEFINITION	Rattus norvegicus clone CH230-53024, WORKING DRAFT SEQUENCE.	
ACCESSION	AC102990	
VERSION	AC102990.6 GI:30521867	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.	
SOURCE	Rattus norvegicus (Norway rat)	
ORGANISM	Rattus norvegicus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.	
REFERENCE	1 (bases 1 to 241724)	
AUTHORS	Muzny,D.,Marle,M.,Metzker,M.,Lee,A.,Abramzon,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Alsbrooks,S.,Amlin,A.,Angulano,D.,Anyalbechi,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,Bryan,N.,Buhay,C.,Burck,P.,Burrell,K.,Calderon,E.,Cardenas,V.,Carter,K.,Cavazos,J.,Ceasar,H.,Chen,X.,Chu,J.,Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Eaves,K.,Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garner,T.,Garza,M.,Gnaniyegis,E.,Geer,K.,Gill,R.,Gard,A.,Guerra,W.,Guerra,W.,Gubarene,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,Harvey,Y.,Havlik,P.,Hawes,A.,Henderson,N.,Hernandez,J.,	

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Hernandez,R., Hines,S., Halam,S.L., Hodson,A., Hogues,M.,
Hollins,B., Howell,S., Huylk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivett,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C.,
Kowitz,C., Kraft,C.L., Lebow,H., Levay,J., Lewis,L., Li,Z., Liu,J.,
Liu,Y., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,D.,
Lorenshewal,L., Louisedge,H., Lozano,R.J., Lu,X., Ma,J.,
Maneshwari,M., Mahindaratne,M., Mahmoud,M., Mallory,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., McNeill,T.Z., Meenen,E.,
Minahney,S., Mclendon,M.P., Monteill,E., Montemayor,J., Moore,S.,
Milosavljevic,A., Miner,G., Minja,E., Mundasa,M., Murphy,M., Nair,L.,
Morjan,M., Morris,K., Norris,S., Nundisa,M., Norris,S., Parks,K.,
Nankervils,C., Neal,D., Newton,N., Nguyen,A., Pal,S., Parks,K.,
Naackelemehe,O., Okwuonu,G., Olapkinsagoon,A., Pali,S., Parks,K.,
Pasternak,S., Paul,H.E., Perez,A., Perez,L., Plannkoeh,C.,
Pioppoer,F., Polndexter,A., Popovic,D., Primus,E., Pu,L., L.,
Piazo,M., Quirroz,J., Rachlin,E., Reeves,K., Regier,M.S., Riggs,F.,
Reilly,B., Reilly,M., Ren,Y., Renter,M., Rose,R., Ruiz,S.J.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,C., Shatsman,S., Shen,H.,
Sanders,W., Savary,G., Scherer,S., Scott,C., Sitter,C.D., Smajz,D.,
Shetty,J., Shvartsbeyn,A., Slisson,I., Silter,C.D., Smajz,D.,
Sneed,A.M., Sodergren,E., Song,X.-B., Sorrelle,R., Sosa,J., Taylor,C.,
Steinle,M., Strong,R., Sutton,A., Syatck,A., Tabot,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tinley,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villason,D., Waldron,L., Walker,B., Wang,J.,
Wang,O., Wang,S., Warren,J., Wleczyk,R., Wooden,H., Worley,K.,
Williams,G., Willison,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yuf,D., Wright,R., Wu,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niefehaenger,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Meinstock,G. and Glbbys,R.A.

COMMENT

Unpublished
Direct Submission
2 (bases 1 to 241724)
Worley,K.C.
Submitted Submission
Directed (24-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 241724)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:22855562.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing results assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly ('a' config-scaffold'). Within each config-scaffold,
individual sequence configs are ordered and oriented, and separate
by sized gaps filled withNs to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
configs within a config-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only configs will be indicated in the feature
table.

----- Genomae Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Project project name: GHEI
Center Clone name: CH230-593024
----- Summary Statistics -----
Assembly program: Atlas 3.0:
Consensus quality: 233518 bases at least Q40
Consensus quality: 235377 bases at least Q30
Consensus quality: 236878 bases at least Q20
Estimated insert size: 245960; sum-of-configs estimation
Quality coverage: 8x in Q20 bases; sum-of-configs estimation

*	30902	31001: gap of unknown length
*	31002	32159: contig of 1158 bp in length
*	32160	32259: gap of unknown length
*	32260	33788: contig of 1529 bp in length
*	33789	33888: gap of unknown length
*	33889	35134: contig of 1246 bp in length
*	35135	35234: gap of unknown length
*	35235	36486: contig of 1252 bp in length
*	36487	36586: gap of unknown length
*	36587	38385: contig of 1799 bp in length
*	38386	38485: gap of unknown length
*	38486	39714: contig of 1229 bp in length
*	39715	39814: gap of unknown length
*	39815	41078: contig of 1264 bp in length
*	41079	41178: gap of unknown length
*	41179	42273: contig of 1095 bp in length
*	42274	42373: gap of unknown length
*	42374	44123: contig of 1750 bp in length
*	44124	44223: gap of unknown length
*	44224	45308: contig of 1685 bp in length
*	45909	46008: gap of unknown length
*	46009	47545: contig of 1537 bp in length
*	47546	47645: gap of unknown length
*	47646	50012: contig of 2367 bp in length
*	50013	50112: gap of unknown length
*	50113	51978: contig of 1866 bp in length
*	51979	52078: gap of unknown length
*	52079	54113: contig of 2035 bp in length
*	54114	54213: gap of unknown length
*	54214	55586: contig of 1373 bp in length
*	55587	55686: gap of unknown length
*	55687	57405: contig of 1719 bp in length
*	57406	57505: gap of unknown length
*	57506	58790: contig of 1285 bp in length
*	58791	58890: gap of unknown length
*	58891	60453: contig of 1563 bp in length
*	60454	60553: gap of unknown length
*	60554	62149: contig of 1596 bp in length
*	62150	62249: gap of unknown length
*	62250	63910: contig of 1661 bp in length
*	63911	64010: gap of unknown length
*	64011	65910: contig of 1900 bp in length
*	65911	66010: gap of unknown length
*	66011	67754: contig of 1744 bp in length
*	67755	67854: gap of unknown length
*	67855	69858: contig of 2005 bp in length
*	69859	69959: gap of unknown length
*	69960	71962: contig of 2003 bp in length
*	71963	72062: gap of unknown length
*	72063	74371: contig of 2309 bp in length
*	74372	74471: gap of unknown length
*	74472	76499: contig of 2028 bp in length
*	76499	76599: gap of unknown length
*	76500	78329: contig of 1730 bp in length
*	78330	78429: gap of unknown length
*	78430	81033: contig of 2604 bp in length
*	81034	81133: gap of unknown length
*	81134	82450: contig of 1317 bp in length
*	82451	82550: gap of unknown length
*	82551	85980: contig of 3330 bp in length
*	85981	85980: gap of unknown length
*	85981	89959: contig of 3979 bp in length
*	89960	90059: gap of unknown length
*	90060	92499: contig of 2440 bp in length
*	92500	92599: gap of unknown length
*	92600	95372: contig of 2773 bp in length
*	95373	95472: gap of unknown length
*	95473	98796: contig of 3324 bp in length
*	98897	98896: gap of unknown length
*	101619	101619: contig of 2723 bp in length
*	101620	101719: gap of unknown length
*	101720	105463: contig of 3744 bp in length
*	105464	105563: gap of unknown length

*	105564	108520: contig of 2957 bp in length
*	108521	108620: gap of unknown length
*	108621	111714: contig of 3094 bp in length
*	111715	111814: gap of unknown length
*	111815	115701: contig of 3887 bp in length
*	115702	115801: gap of unknown length
*	115802	117720: contig of 1919 bp in length
*	117721	117820: gap of unknown length
*	117821	121379: contig of 3459 bp in length
*	121380	121379: gap of unknown length
*	121380	125489: contig of 4110 bp in length
*	125490	125899: gap of unknown length
*	125899	128897: contig of 3308 bp in length
*	128898	133134: gap of unknown length
*	133135	133234: contig of 4137 bp in length
*	133235	133234: gap of unknown length
*	137403	137402: contig of 4168 bp in length
*	137503	137502: gap of unknown length
*	142643	142642: contig of 5140 bp in length
*	142743	142742: gap of unknown length
*	148553	148552: contig of 5810 bp in length
*	148553	148552: gap of unknown length
*	158806	158806: contig of 10154 bp in length.
FEATURES		
source		
1. .158806		
Location/Qualifiers		
/organism="Mus musculus"		
/mol_type="genomic DNA"		
/db_xref="taxon:10090"		
/chromosome="Y/NK"		
/clone="RP24-335114"		
1. .1262		
/note="assembly_name:Contig10"		
misc_feature		
1363. .2809		
/note="assembly_name:Contig22"		
misc_feature		
2910. .4113		
/note="assembly_name:Contig34"		
Query Match		
Best Local Similarity 61.1%: Score 229: DB 2: Length 158806:		
Matches 312: Conservative 0: Mismatches 50: Indels 13: Gaps 4:		
QY	6	CCCCGACCCCTCATCTCGGCGCCGCCCTTCCTCCACCCCTCCGAGCCCTAAAGGG 65
DB	138522	CCCGAATCTTTCATGCGGGCTGCAGCATCTCTCCACCCCTCAGAGAGCGAA----- 138576
QY	66	CGGCGGGGCGCCAGCCGAGGCGCTGCGCTGACCCCGAGCGGAAAGGCCCACTTAGG 125
DB	138577	---GGACAGCGGAGCGAGGGGCGAGCGCGTGACCCGAGGGAAGGCGCCAGTTCGG 138633
QY	126	TCTTAATGCGGGTGGCTCTCTTGTGACAGGCGGCTTTGGGGACAACAGCGGGAGAG 185
DB	138634	TCTTAATGCGGGTGGCTCTCTTGTGACAGGCGGCTTTGGGGACAACAGCGGGAGAG 138693
QY	186	AGATAAGGTGACATACACAGACAGATTGGTGGCGGCTGATACCTCTCCGACAGG 245
DB	138694	AGATAAGGTGACATACACAGACAGATTGGTGGCGGCTGATACCTCTCCGACAGG 138753
QY	246	AAACGCGAGCTATT--AAAGACCTATGCACTTATCTTCC--TGAAAGCT 301
DB	138754	AAATTCGGAGCTATTAAAGGCGCTATGCACTTATCTTCCCGAGGAAACTT 138813
QY	302	CTTGCGGAGGACAAAGATGT-TCCCGGCTAAAGACAAAGGCGACACAGAGAGGT 360
DB	138814	CTTGCGGAGGACAAAGATGTCTCCCTTAAAGATACAAAGGCGACACAGAGAGGT 138873
QY	361	CTGCACAGCGACGC 375
DB	138874	TTGTACAGCGACGC 138888
RESULT 6		
CNS08CBD/c		
LOCUS	CNS08CBD	211509 bp DNA linear PRI 05-SEP-2002

[illegible][illegible]

```

DEFINITION Homo sapiens zinc finger transcriptional regulator (GOS24) gene,
complete cds.
ACCESSION M92844
VERSION M92844.1 GI:2072389
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3889)
AUTHORS Blum, S., Forsdyke, R.E. and Forsdyke, D.R.
TITLE Three human homologs of a murine gene encoding an inhibitor of stem
cell proliferation
JOURNAL DNA Cell Biol. 9 (8), 589-602 (1990)
MEDLINE 91103879
PUBMED 2271120
REFERENCE 2 (sites)
AUTHORS Taylor, G.A., Lal, W.S., Oakley, R.J., Seidlin, M.F., Shows, T.B.,
Eddy, R.L. Jr. and Blackshear, P.J.
TITLE The human TTP protein: sequence, alignment with related proteins,
and chromosomal localization of the mouse and human genes
JOURNAL Nucleic Acids Res. 19 (12), 3454 (1991)
MEDLINE 91288233
PUBMED 2062660
REFERENCE 3 (bases 1 to 3889)
AUTHORS Heximer, S.P. and Forsdyke, D.R.
TITLE A human putative lymphocyte G0/G1 switch gene homologous to a
rodent gene encoding a zinc-binding potential transcription factor
JOURNAL DNA Cell Biol. 12 (1), 73-88 (1993)
MEDLINE 8422274
PUBMED
REFERENCE 4 (bases 1 to 3889)
AUTHORS Heximer, S.P., Cristillo, A.D., Russell, L. and Forsdyke, D.R.
TITLE RT-PCR analysis of RNA of the CCH zinc finger protein encoding
gene GOS24 (TIS11/TTP/NUP475) in cultured human blood mononuclear
cells
JOURNAL Unpublished
COMMENT On May 6, 1997 this sequence version replaced gi:183444.
FEATURES
SOURCE
location/Qualifiers
1..3889
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="19q13.1"
/tissue_type="placenta"
549..630
/feature="GOS24"
exon
/number=1
/evidence=experimental
join(607..630,1441..2397)
/feature="GOS24"
CDS
join(607..630,1441..2397)
/feature="GOS24"
exon
/codon_start=1
/product="zinc finger transcriptional regulator"
/protein_id="AAC37600.1"
/db_xref="GI:183445"
/translation="MDLTALYESLISLSPDPVPSDHGTESSFGWSSGFGWLSLSPD
SSPGVTSRLPGSTSLVEGRSCGWPPPGFAPLAPRLGTELSPPSPATSTPS
RYKELCRTESEGRCHRYGAKCGAFHAGELROANRHPKYTELCHEKFFYLOGRQYGS
RCHFTIHNPSEDLAPGHPPIYROSITSPSGILSPGRTSPPPGLAGLSSSSFSFSS
PPPPGDLPLSFAFSAAGPTPLARDPPTVCCPCSCRRATPTISWAGPLGGIYRTSYVS
LGSPPDEYASSGSSLSGSDSPFEAGVFAFPQVPAAPARPLIFNRISVSE"
1441..3103
/feature="GOS24"
exon
/number=2
/evidence=experimental
2978..2986
/misc_feature
/note="TA-rich conserved element (TARCE)"
3163..3168
/misc_feature
/note="TG-rich RNA polymerase II termination element"
3198..3889
/misc_feature
/note="Cpg island"

```

```

BASE COUNT 648 a 1304 c 1078 g 850 t 9 others
ORIGIN chromosome 19q13.1.
Query Match 11.7%; Score 43.8; DB 9; Length 3889;
Best Local Similarity 59.1%; Pred. No. 0.73;
Matches 75; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 2 GGGCCCCCGCCACATCCGCTCCGCGCCCTTCTCTCCACCCCTCCGACCCCTTAA 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GGGCAGGCGTCCCGCATCCGACGCCGCCCTTCCGACCATTCGCCGCTGGTAC 360
62 GGGCGCGCGGGGCCCAAGCGCGCTGCGCTGACCCCGAGGGAAGGCCCCCATGC 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GCGTGTCCACCGGCAAGCTCAGGCGCTCCAGCGCGCGGAGGGAACCACTGC 420
QY 122 TAGGTTC 128
||| |||
Db 421 CAGGGCC 427

```

```

RESULT 8
AC058782 171852 bp DNA linear HTG 22-FEB-2001
LOCUS Homo sapiens chromosome 1 clone RP11-285F7, WORKING DRAFT SEQUENCE,
8 unordered pieces.
DEFINITION AC058782
ACCESSION AC058782.11 GI:12331477
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 171852)
AUTHORS Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D.,
Fedorpiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R.,
Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J.,
Ramirez, D., Wilhelm, J., Yu, S. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT On Jan 22, 2001 this sequence version replaced gi:12203632.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development
Center
Center code: SDSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: RP11-285F7
Center clone name: RP11-285F7
----- Summary Statistics
Sequencing Vector: M13mp18, X02513
Chemistry: Dye-terminator 08 of reads
Chemistry: Dye-terminator Big dye; 1008 of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 165994 bases at least Q40
Consensus quality: 167134 bases at least Q30
Consensus quality: 167611 bases at least Q20
Insert size: 165696; agarose-fp
Insert size: 171152; sum-of-contigs
Quality coverage: 9.4x in Q20 bases; agarose-fp
Quality coverage: 9.1x in Q20 bases; sum-of-contigs.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```



```
* 88572 88671: gap of 100 bp
* 88672 102917: contig of 14246 bp in length
* 102918 103017: gap of 100 bp
* 103018 120081: contig of 17064 bp in length
* 120082 120181: gap of 100 bp
* 120182 136057: contig of 15876 bp in length
* 136058 136157: gap of 100 bp
* 136158 156302: contig of 20145 bp in length
* 156303 156402: gap of 100 bp
* 156403 167887: contig of 11385 bp in length
* 167888 167888: gap of 100 bp
* 167888 183109: contig of 15222 bp in length
* 183110 183209: gap of 100 bp
* 183210 196606: contig of 13397 bp in length
* 196607 196706: gap of 100 bp
* 196707 218924: contig of 22218 bp in length
* 218925 219024: gap of 100 bp
* 219025 249572: contig of 30548 bp in length
* 249573 249672: gap of 100 bp
* 249673 287188: contig of 37516 bp in length
* 287189 287288: gap of 100 bp
* 287289 289388: contig of 2100 bp in length
* 289389 289488: gap of 100 bp
* 289489 291859: contig of 2371 bp in length
* 291860 291959: gap of 100 bp
* 291960 294128: contig of 2169 bp in length
* 294129 296882: gap of 100 bp
* 296883 296982: gap of 100 bp
* 296983 299603: contig of 2621 bp in length
* 299604 301793: gap of 100 bp
* 301794 301893: gap of 100 bp
* 301894 305076: contig of 3183 bp in length
* 305077 305176: gap of 100 bp
* 305177 308174: contig of 2998 bp in length
* 308175 308274: gap of 100 bp
* 308275 310542: contig of 2268 bp in length
* 310543 310642: gap of 100 bp
* 310643 313870: contig of 3228 bp in length
* 313871 313970: gap of 100 bp
* 313971 317124: contig of 3154 bp in length
* 317125 317224: gap of 100 bp
* 317225 320595: contig of 3371 bp in length
* 320596 320695: gap of 100 bp
* 320696 323155: contig of 2460 bp in length
* 323156 323255: gap of 100 bp
* 323256 326284: contig of 3029 bp in length
* 326285 326384: gap of 100 bp
* 326385 329626: contig of 3242 bp in length
* 329627 329726: gap of 100 bp
* 329727 333482: contig of 3756 bp in length
* 333483 333582: gap of 100 bp
* 333583 337189: contig of 3607 bp in length
* 337190 337289: gap of 100 bp
* 337290 341875: contig of 4586 bp in length
* 341876 341975: gap of 100 bp
* 341976 344388: contig of 2413 bp in length
* 344389 344488: gap of 100 bp
* 344489 349154: contig of 4666 bp in length
* 349155 349254: gap of 100 bp
* 349255 353824: contig of 4570 bp in length
* 353825 353924: gap of 100 bp
* 353925 357953: contig of 4029 bp in length
* 357954 358053: gap of 100 bp
* 358054 363695: contig of 5642 bp in length
* 363696 363795: gap of 100 bp
* 363796 367278: contig of 3483 bp in length
* 367279 367378: gap of 100 bp
* 367379 371902: contig of 4524 bp in length
* 371903 372002: gap of 100 bp
* 372003 377117: contig of 5115 bp in length
* 377118 377217: gap of 100 bp
```

```
* 377218 382171: contig of 4954 bp in length
* 382172 382271: gap of 100 bp
* 382272 388633: contig of 6562 bp in length
* 388634 388933: gap of 100 bp
* 388934 393530: contig of 4597 bp in length
* 393531 393630: gap of 100 bp
* 393631 398934: contig of 5304 bp in length
* 398935 399034: gap of 100 bp
* 399035 404551: contig of 5517 bp in length
* 404552 404651: gap of 100 bp
* 404652 410493: contig of 5842 bp in length
* 410494 410593: gap of 100 bp
* 410594 415944: contig of 5351 bp in length
* 415945 416044: gap of 100 bp
* 416045 424416: contig of 8372 bp in length
* 424417 424516: gap of 100 bp
* 424517 430287: contig of 5771 bp in length
* 430288 430387: gap of 100 bp
* 430388 437707: contig of 7320 bp in length
* 437708 437807: gap of 100 bp
* 437808 445871: contig of 8064 bp in length
* 445872 445971: gap of 100 bp
* 445972 455450: contig of 9479 bp in length.
* 455450: Location/Qualifiers
  1. .455450
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="11"
    /map="11"
    /clone="MGS1-213E3"
    /clone_1lb="Genome Systems Mouse BAC 1"
    /clone_1lb="119379 a 104363 c 107639 g 117130 t 6939 others"
```

BASE COUNT 119379 a 104363 c 107639 g 117130 t 6939 others

ORIGIN

Query Match 11.2% Score 42; DB 2; Length 110000;
Best Local Similarity 56.5%; Pred. No. 3.5;
Matches 78; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

```
QY 5 CCCCCGACCCCTATCTGCTCCCGCCCTTCTCTGCACCCCTCCGACCCCTAAAGG 64
   || || || || || || || || || || || || || || || || || || || ||
Db 72293 CTTCCCTCCCTCTCTCTCCGCGCGCGCCCTCCCTCCCTCCCGGGGCGGAGCG 72352

QY 65 GCGCGGGGGCCCAAGCGAGGCGCGCTGACCCAGCGAGCGGAGCGCCAGTCTAG 124
   || || || || || || || || || || || || || || || || || || || ||
Db 72353 GGGGGAGGAGCGCGCGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 72412

QY 125 GTCTATGCGGCGTGGCG 142
   || || || || || || || || || || || || || || || || || || || ||
Db 72413 GCGCGGGGGGGGGGGCG 72430
```

RESULT 10

AC026512/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AUTHORS

AUTHORS

AUTHORS

AUTHORS

AUTHORS

AUTHORS


```

repeat_region /rpt_family="L2"
complement(4926..5672)
/rpt_family="L2"
repeat_region complement(6570..6829)
/rpt_family="MIR"
repeat_region 7457..7493
/rpt_family="(TTTAA)n"
complement(7884..8219)
/rpt_family="L2"
repeat_region 8517..8554
/rpt_family="(CA)n"
repeat_region 8996..9563
/rpt_family="MER90"
repeat_region 9649..9718
/rpt_family="MER5B"
repeat_region 10054..10253
/rpt_family="L2"
repeat_region 10325..10440
/rpt_family="L2"
repeat_region 10610..10822
/rpt_family="MIR"
complement(10881..11030)
/rpt_family="MER5A"
repeat_region 11046..11146
/rpt_family="L2"
repeat_region 11983..12096
/rpt_family="MER5A"
complement(12275..12452)
/rpt_family="MER5A"
complement(13065..13238)
/rpt_family="Charlie8"
repeat_region 13749..13788
/rpt_family="MLT1H1"
repeat_region 13789..14092
/rpt_family="Alusp"

```

Query Match 11.1%; Score 41.8; DB 9; Length 90707;

Best Local Similarity 49.3%; Pred. No. 3.9; Mismatches 112; Indels 0; Gaps 0;

Matches 109; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

Db 10 GCACCCCTCATCTGCGTCCGCCCTTCTCTCCACCTCCCGGACCCCTAAAGGGCGGC 69
68040 GCGCCCGCGCGGATCCACCTGCACCGCCTCTCTGCGCCCGGAGCTCCAAACCCGCG 67981
OY 70 GGGGCCCAAGCCGAGGGCGCTGCGCCGACCCCGAGGGGAAAGGGCCAGTCTAGTCT 129
Db 67980 AAGCTTGGCGCCACAGGAGCGCGCGGACGCCGCCAGAGCCGCCCGCTCAGAGAGCC 67921
OY 130 AATCGGGGTGCGTCTCTCTTGGACAGCGCGGCTTGGGGACACAGCGGGGAGAGAGAT 189
Db 67920 GCGTCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 67861
OY 190 AAGGTGACATACAGACAGATTTGGTGGCGGCGCTGATC 230
Db 67860 GCGCTTCTTACCTGACGACGCTGCGAGGGGAGACGACAC 67820

```

RESULT 11

AL773527 137286 bp DNA linear MAM 07-SEP-2002
 LOCUS Pig DNA sequence from clone XX-707FL, complete sequence.
 DEFINITION AL773527
 ACCESSION AL773527.8 GI:21998261
 VERSION
 KEYWORDS HTG.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 137286)
 AUTHORS Sehra,H.
 TITLE Direct Submission
 JOURNAL Submitted (06-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

COMMENT

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Jul 26, 2002 this sequence version replaced gi:21912758.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Sw.,
 SWISSPROT; Tr., TrEMBL; Wp., WormPeP; Information on the WormPeP
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep XX-707FL is from
 a pig library
 VECTOR: pBelBAC11.
 Location/Qualifiers

FEATURES

source
 1..137286
 /organism="Sus scrofa"
 /mol_type="genomic DNA"
 /db_xref="taxon:9823"
 /clone="XX-707FL"
 /clone_lib="SBAB"

BASE COUNT 32551 a 35854 c 35765 g 33116 t

ORIGIN

Query Match 11.1%; Score 41.6; DB 4; Length 137286;

Best Local Similarity 50.5%; Pred. No. 4.7; Mismatches 99; Indels 0; Gaps 0;

Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

```

OY 25 CTCGGCGCCCTTCTCTCACCTTCCGACCCCTTAAGGGGCGGGGCCCAAGCCGAG 84
Db 63759 CACACGGCAGCAGCGCGCCCGCCCTTACGCCACGACGCGCGCGCGC 63818
OY 85 GCGCTGCGCTTACCCCGAGCGAGGGCCCACTAGTCTCTAATGGCGGTGGCGTC 144
Db 63819 GCGCGCGCCCGTGCACCCCGCGCTGCGCTCCCAAGGCTTGGCTGTGGGGC 63878
OY 145 TCTTTTGCACAGCGCGGTTTGGGACACACGCGGAGAGAGATTAAGTACATACAG 204
Db 63879 GCAGCGGCGCCCGGGGGGGTGGGAAATAGGGGGTTAAAAAAGCGCGGACCGG 63938
OY 205 AGCAGATTGGTGGCGCGCGC 224
Db 63939 GCGAGGTGAGAGCGCGCTGC 63958

```

RESULT 12

OSJN00209/c 164907 bp DNA linear PLN 17-JUL-2002
 LOCUS Oryza sativa genomic DNA, chromosome 4, BAC clone: OSJNB0058J09,
 DEFINITION complete sequence.
 ACCESSION AL663008
 VERSION AL663008.2 GI:21742449
 KEYWORDS HTG.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;
 Euphorbiales; Oryzaceae; Oryza.
 REFERENCE 1

/gene="B1144G04.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90016.1"
/db_xref="GI:20161086"
/translation="MERSVEKAPASALSWGLYSENFIDPSCVPPRASDDSESP
DASFPDPTCVLSSTSDRISVLPGRCPAGVGSGPPSSPAFLGLPARSSGILP
PAGVHGAK"
complement(join(32456..32459,32541..32698))
/gene="B1144G04.8"
complement(join(32456..32459,32541..32698))
/gene="B1144G04.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB90017.1"
/db_xref="GI:20161087"
/translation="MPPIIPEPFRAKONGKCRDNTWKCRLQLQELKAQEKPEAHIAKN
PRISDGRGM"
/join(33277..33381,33495..33649,33772..33817,33914..34005,
34106..34202,34291..34386,34480..34577,35339..35936)
/gene="B1144G04.9"
/join(33277..33381,33495..33649,33772..33817,33914..34005,
34106..34202,34291..34386,34480..34577,35339..35936)
/gene="B1144G04.9"
/note="hypothetical protein
similar to receptor protein kinase"
/codon_start=1
/protein_id="BAB90018.2"
/db_xref="GI:24636778"
/translation="MCELRISRRRTTAGLGLYLILGYKAAVPEYFVLLINYM
SVYIIPRTSNLWVRLVREWOFICIGWNPRIPEASLILPYPLKSTWELAMP
VIMRADDPDNPVFLVREWOFICIGWNPRIPEASLILPYPLKSTWELAMP
PIYSYVENPCCQCDTQDMADAKGLVSDHMGVRSQSPQLILVLYNP
SKVSTAAVASSPLPESSLSLSPSKRRYHGERKORRLITGKRPRDPE
LNCTSLQFLDMSTVNGNMQILGKFTLRVLHNNHNGRTYSEVLRPLPAR
LNISFOYGELEPLVADAMKSLKPYMLLNNSGIPATYPPNFYTYWTRNCSL
WRLLPHHRHGRVHNGDPASEVFGSGPLGPRVALL"
/join(37887..37936,38063..38177,38961..39149,40379..40738,
40819..40882,41046..41185,41291..41599,41816..42070)
/gene="B1144G04.10"
/join(37887..37936,38063..38177,38961..39149,40379..40738,
40819..40882,41046..41185,41291..41599,41816..42070)
/gene="B1144G04.10"
/codon_start=1
/product="putative ubiquitin-conjugating protein"
/protein_id="BAB90019.1"
/db_xref="GI:20161089"
/translation="MPSPPVPSSTAEGSSKNSPTKSLSLGLDAVEVLPASATAGM
Query Match 10.8% Score 40.4; DB 8; Length 194459;
Best Local Similarity 51.7%; Pred. No. 10;
Matches 92; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
HTG: HTGS_PHASE1: HTGS_DRAFT.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 200557)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 200557)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (01-OCT-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BB0323105
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 190066 bases at least Q40
Consensus quality: 192485 bases at least Q30
Consensus quality: 194040 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1128: contig of 1128 bp in length
* 1129 1228: gap of unknown length
* 1229 2347: contig of 1119 bp in length
* 2348 2447: gap of unknown length
* 2448 3450: contig of 1003 bp in length
* 3451 3550: gap of unknown length
* 3551 4960: contig of 1410 bp in length
* 4961 5060: gap of unknown length
* 5061 6079: contig of 1019 bp in length
* 6080 6179: gap of unknown length
* 6180 7276: contig of 1097 bp in length
* 7277 7375: gap of unknown length
* 7377 8642: contig of 1266 bp in length
* 8643 8743: gap of unknown length
* 8743 10002: contig of 1260 bp in length
* 10003 10102: gap of unknown length
* 10103 11298: contig of 1196 bp in length
* 11299 11398: gap of unknown length
* 11399 13086: contig of 1688 bp in length
* 13087 13187: gap of unknown length
* 13188 16666: contig of 3480 bp in length
* 16667 16767: gap of unknown length
* 16768 20038: contig of 3272 bp in length
* 20039 20138: gap of unknown length
* 20139 24119: contig of 3981 bp in length
* 24120 24220: gap of unknown length
* 24221 30932: contig of 6713 bp in length
* 30933 31032: gap of unknown length
* 31033 35374: contig of 4342 bp in length
* 35375 35475: gap of unknown length
* 35476 41696: contig of 6222 bp in length
* 41697 41796: gap of unknown length

